The role of human CBX proteins in human benign and malignant hematopoiesis

Jung, Johannes

4 THE ROLE OF CBX PROTEINS IN HUMAN BENIGN AND MALIGNANT HEMATOPOIESIS

Jung J[1], Buisman SC[1], Weersing E[1] ,
Dethmers-Ausema B[1], Zwart E[1], Schepers H[2],
Dekker MR[3], Lazare SS[1], Hammerl F[1], Paeschke K[1],
Juranek S[1], Skokova J[4], Kooistra SM[5],
Klauke K[1], Poot RA[3], Bystrykh LV[1],* and de Haan G[1].*

Affiliations:

1. European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, the Netherlands
2. Department of Hematology, University Medical Center Groningen, University of Groningen, the Netherlands
3. Department of Cell Biology, ErasmusMC, Rotterdam, the Netherlands
4. University Hospital of Tübingen, Germany
5. Department of Neuroscience, Section Medical Physiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

*Corresponding authors: (l.bystrykh@umcg.nl and g.de.haan@umcg.nl).

In preparation

SUMMARY:

In this study we demonstrate that among all five CBX Polycomb proteins, only CBX7 possesses the ability to control self-renewal of human hematopoietic stem- and progenitor cells (HSPCs). Xenotransplantation of *CBX7*-overexpressing HSPCs resulted in increased multi-lineage long-term engraftment and myelopoiesis. Gene expression and chromatin analyses revealed perturbations in genes involved in differentiation, DNA and chromatin maintenance, and cell cycle control. *CBX7* is upregulated in AML and its genetic or pharmacological repression in AML cells inhibited proliferation and induced differentiation. Mass spectrometry analysis revealed novel non-histone protein interactions between CBX7 and the H3K9 methyltransferases SETDB1 and EHMT1 and -2. These CBX7-binding proteins possess a trimethylated lysine peptide motif highly similar to the canonical CBX7 target H3K27me3. Depletion of SETDB1 in AML cells phenocopied repression of CBX7. We identify CBX7 as an important regulator of self-renewal and uncover novel, non-canonical crosstalk between epigenetic pathways revealing new therapeutic opportunities for leukemia.

SIGNIFICANCE:

Epigenetic modifiers are important regulators of hematopoietic stem cells and are frequently mutated or aberrantly expressed in hematological malignancies. Because epigenetic changes are in general reversible, they represent putative druggable targets. Here we show that CBX7, known to direct the Polycomb Repressive Complex-1 to loci marked by H3K27me3, regulates self-renewal of benign hematopoietic stem as well as leukemic cells. We demonstrate that CBX7 does not only bind to H3K27me3, but also interacts with multiple non-histone proteins, including enzymes involved in H3K9 methylation, which themselves harbor trimethylated lysine residues. This non-canonical role for CBX7 mediates crosstalk between Polycomb activity and other pathways involved in epigenetic repression. Disruption of canonical and non-canonical interactions leads to differentiation of leukemic cells, and thus suggests novel therapeutic opportunities.

INTRODUCTION

Hematopoietic stem cells (HSCs) are able to self-renew and differentiate into all mature blood cells to ensure peripheral blood cell homeostasis during adult lifespan. In these primitive cells, the choice between self-renewal and differentiation must be well balanced to avoid either cytopenia or hyperproliferative conditions like leukemia. Self-renewal and differentiation are accompanied and controlled by a multitude of epigenetic changes of DNA and of histone proteins (Kamminga et al., 2006; Klauke et al., 2013; Rizo et al., 2008; Tadokoro et al., 2007). One important family of epigenetic regulators that is critical for stem cells is represented by the Polycomb group (PcG) genes.

PcG genes encode for chromatin-associated proteins, which assemble in various multimeric protein complexes and contribute to the regulation of gene expression patterns by posttranslational modifications of histone tails (Bracken et al., 2006; Cao et al., 2005).

The two best-characterized PcG protein complexes are the canonical Polycomb Repressive Complex 1 (PRC1) and -2 (PRC2). The canonical PRC1 is characterized by the presence of at least one of the five Polycomb chromobox proteins (CBX2, 4, 6, 7 and 8). Many functional and molecular studies have shown similar and overlapping binding patterns of PRC1- and PRC2-protein containing complexes (Comet and Helin, 2014; Morey et al., 2012). Although the enzymatic activity of many individual epigenetic writers and erasers has been elucidated, our understanding of the biological role and the molecular dynamics of epigenetic protein complexes is still limited.

CBX proteins are characterized as chromodomain-containing proteins, recognizing trimethylated lysine 27 on histone H3 (H3K27me3), which is deposited by EZH1/2 (Fischle et al., 2003; Min et al., 2003). After recognition of H3K27me3 by the CBX proteins, the catalytic subunit of PRC1, RING1A/B, ubiquitinates H2AK119 (Cao et al., 2005) leading to repression of transcription through chromatin compaction and inhibition of RNA Polymerase II (Stock et al., 2007). Beyond this classical PRC2/PRC1 recruitment model, evidence is emerging for a far more diverse and complicated composition and recruitment process. Most notably, it has become apparent that a plethora of distinct PRC1 complexes exist, some of which contain RYBP instead of CBX (Tavares et al., 2012). Furthermore, PRC1 can be present at genomic loci in the absence of any PRC2 activity (Kahn et al., 2016).

Notwithstanding our limited understanding of the complex protein-protein and protein-DNA interactions in which the PcG proteins are involved, it has become evident that PcG proteins are important regulators of self-renewal and differentiation of many types of pluripotent and adult stem cells (Morey et al., 2013). Indeed, deregulation of their expression or mutations in genes coding for PcG proteins can result in cancer development. We have previously shown that overexpression of the H3K27 methyltransferase *Ezh2* in murine HSCs prevents their exhaustion in serial transplantation experiments (Kamminga et al., 2006). Furthermore, both EZH2 and BMI1 are important regulators of self-renewal of normal murine and human hematopoietic stem cells (Rizo et al., 2008). Interestingly, mutations in the *EZH2* gene were later found in patients with myelodysplastic syndromes and acute myeloid leukemia (Cancer Genome Atlas Research et al., 2013; Nikoloski et al., 2010)

More recently, we showed that Cbx7, but not Cbx2, -4, or -8, is a potent regulator of self-renewal of murine hematopoietic stem cells and its enforced overexpression resulted in increased self-renewal and in phenotypically diverse leukemias (Klauke et al., 2013). In human cells, systematic short hairpin-mediated repression of all CBX proteins in $CD34^+$ cord blood cells resulted in decreased proliferation and colony-forming unit ability. In this experiment knockdown of *CBX2* was shown to be most detrimental (van den Boom et al., 2013).

Collectively these studies highlight the relevance of PcG proteins, and particularly CBX proteins, in maintaining blood cell homeostasis. As epigenetic changes are in principle reversible, elucidating the function of epigenetic writers, readers, and erasers in the context of healthy and malignant hematopoiesis is indispensable for identifying novel therapeutic targets.

Therefore, in the current study, we asked to what extent different CBX proteins are able to affect the balance between self-renewal and production of mature blood cells of normal human cord blood-derived primitive $CD34^+$ cells. We identify *CBX7* as a potent inducer of self-renewal. Reversely, repression of *CBX7* in AML cells results in their terminal differentiation. In addition, we identify novel evolutionary conserved non-histone interaction partners of CBX7. These novel interaction partners include multiple epigenetic enzymes, most notably SETDB1, EHMT1, and EHMT2, which are all H3K9 methyltransferases that carry a potential lysine site for trimethylation. These sites are in a conserved peptide context, which is similar to H3K9me3 and H3K27me3. Importantly, depletion of SETDB1, similar to

CBX7, also induced differentiation of AML cells, suggesting that at least part of the self-renewal potential of CBX7 is dependent on its interaction with an H3K9-methyltransferase. H3K27me3 and H3K9me3 ChIP-seq and RNA-seq experiments revealed direct and indirect CBX7 targets which comprise a complex network of both classical histone modifications and novel epigenetic interactions that collectively control the balance between self-renewal and differentiation in primitive human hematopoietic cells.

RESULTS

CBX7 enhances self-renewal of human CD34$^+$ cord blood cells *in vitro*

To assess the role of the five different PRC1-CBX proteins on hematopoietic progenitor function, we overexpressed *CBX2*, *4*, *6*, *7* and *8* in CD34$^+$ cord blood cells and performed colony-forming unit (CFU) assays. Whereas overexpression of *CBX7* and *CBX8* resulted in increased CFU-frequencies, overexpression of *CBX2* and -4 resulted in lower CFU-frequencies in comparison to empty vector control (EV). Overexpression of *CBX6* had no discernable effect (Figure 1A). Although *CBX8* overexpression resulted in a slightly higher CFU frequency in comparison to *CBX7*, replating of *CBX7* overexpressing cells from a primary plate resulted in higher CFU-frequency (Figure 1B). In line with these data, *CBX8* overexpressing CD34$^+$ HSPCs showed no proliferative advantage in comparison to control cells in a cytokine-driven suspension culture, whereas *CBX7* overexpressing CD34$^+$ HSPCs showed a strong proliferative advantage and could be kept in culture up to ten weeks (Figure 1C and Supplementary Figure 1A). To determine the role of the five different CBX proteins in regulating the most primitive cell compartment, we performed cobblestone area-forming cell (CAFC) assays, that were evaluated 35 days after seeding of transduced and sorted CD34$^+$ cells. *CBX7* overexpression increased the CAFC frequency ~10 fold, whereas *CBX8* overexpression resulted in a smaller increase in CAFC frequency (Figure 1D). In contrast, overexpression of *CBX4* decreased the cobblestone area-forming cell frequency dramatically (~50-fold), while overexpression of *CBX2* and *CBX6* had no effect. We next tested whether CBX7 is essential for self-renewal of human primitive CD34$^+$ cells by performing short hairpin

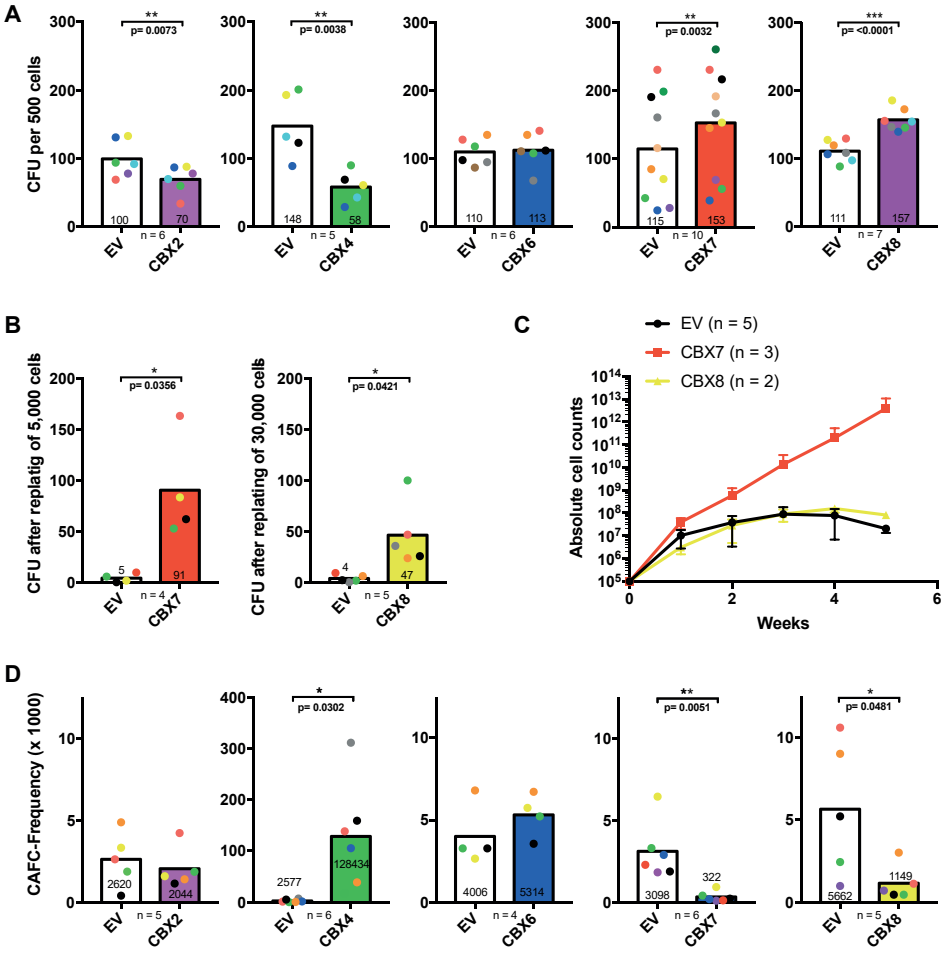
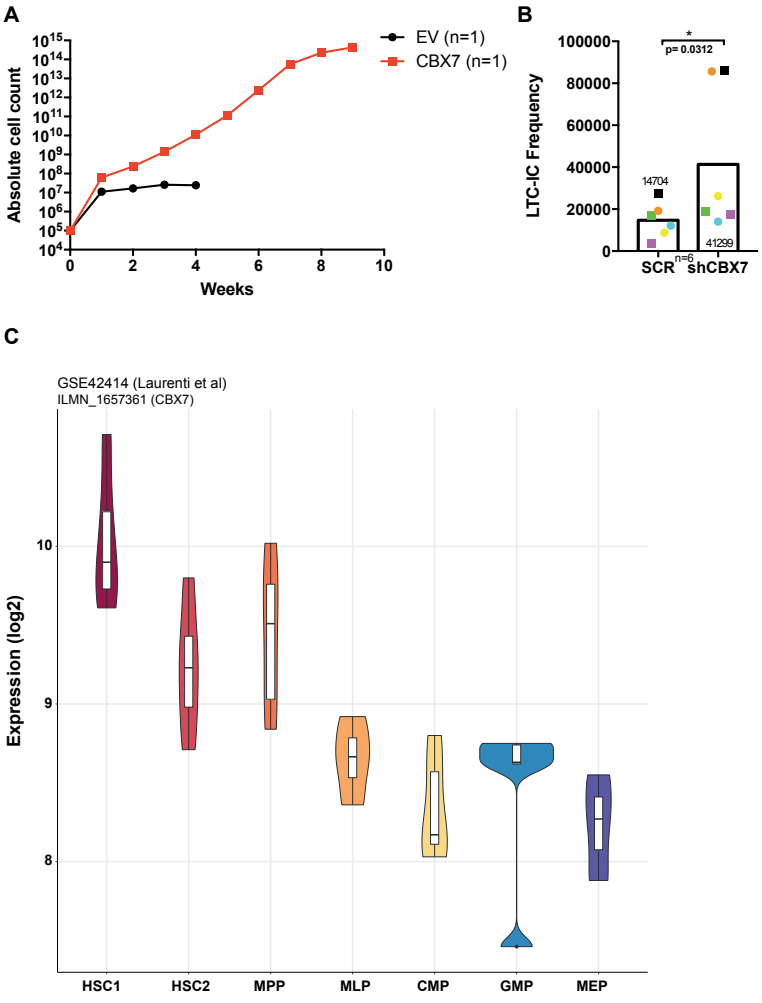


Figure 1:

Enforced retroviral overexpression of *CBX2*, *4*, *6*, *7* or *8* reveals distinct effects on human CD34$^+$ cord blood-derived progenitors and primitive cells in vitro.

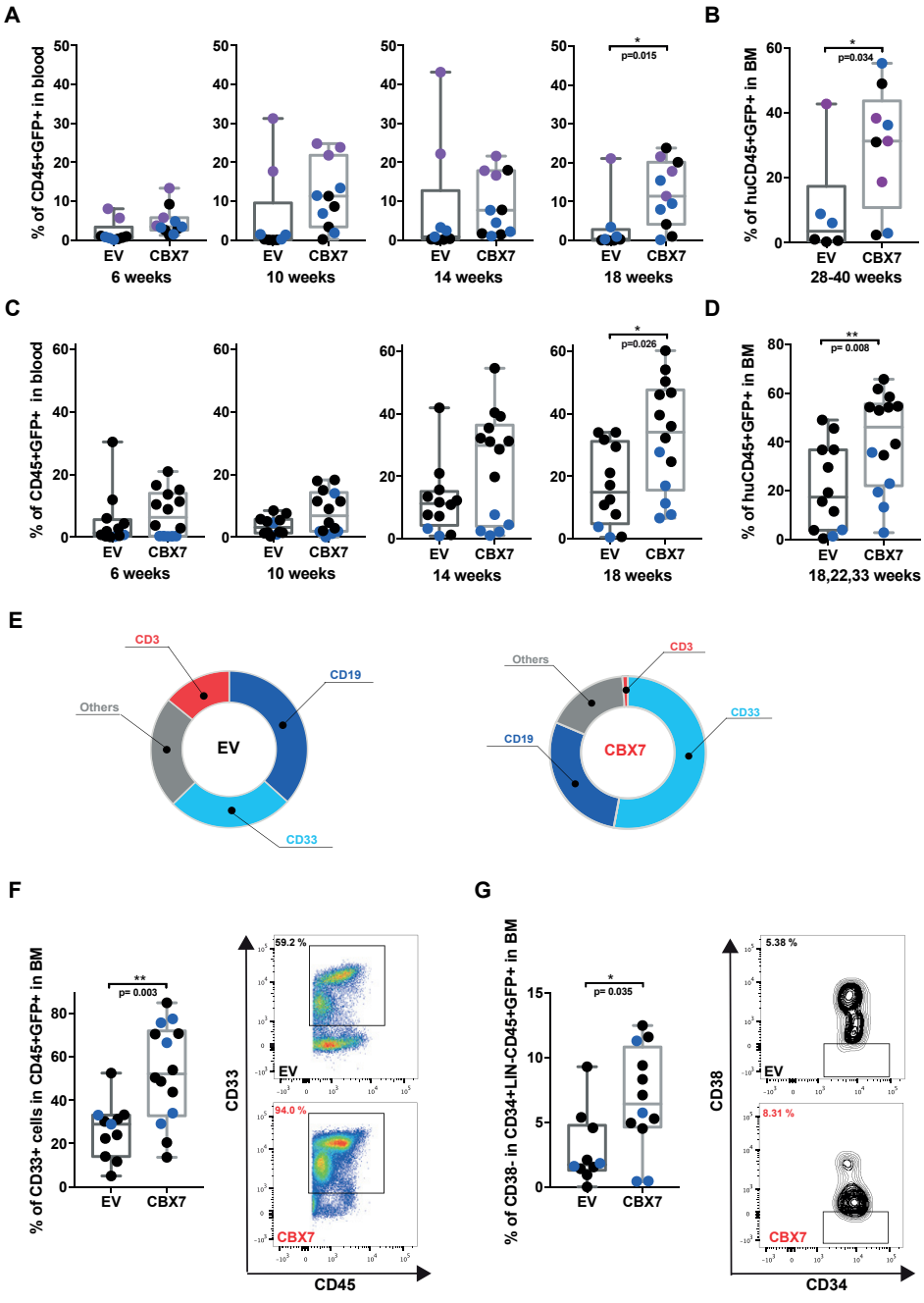
(**A**) CFU-frequencies of cord blood-derived CD34$^+$ cells overexpressing *CBX2*, *4*, *6*, *7*, or *8*. (**B**) CFU-frequencies after replating of 5,000 *CBX7* overexpressing cells or 30,000 *CBX8* overexpressing cells. (**C**) Absolute cell counts in cytokine-driven liquid culture of cord blood-derived CD34$^+$ cells transduced with an empty vector (EV), or upon *CBX7* or *CBX8* overexpression. (**D**) Day 35 cobblestone area-forming cell frequency (CAFC) of cord blood-derived CD34$^+$ cells overexpressing *CBX2*, *4*, *6*, *7*, or *8*. (The Y-axis indicates the number of cells that need to be plated for a CAFC to develop.)

[Each graph represents multiple independent experiments with different cords. Identically colored circles indicate paired experimental and control samples that originate from the same cord. Statistical analysis was performed using a two-tailed paired t-test.]



Supplementary Figure 1, related to Figure 1

(**A**) Absolute cell counts in cytokine-driven liquid culture of cord blood-derived $CD34^+$ cells transduced with an empty vector (*EV*) or *CBX7* overexpressing vector. (**B**) Frequency of long-term culture initiating cells of cord blood-derived $CD34^+$ cells upon knockdown with two different short-hairpins against *CBX7*. The number of cells that need to be plated for one LTC-IC to develop is indicated. Each graph represents multiple independent experiments with different cords. Identically colored circles indicate paired experimental and control samples that originate from the same cord. Rectangles represent knockdown with shCBX7#1, circles represent knockdown with shCBX7#2 (n = 6). Statistical significance was determined using Wilcoxon matched-pairs signed rank test.(**C**) Log2 mRNA expression levels of *CBX7* (ILMN_1657361) in different subsets of sorted $CD34^+$ cord blood derived hematopoietic stem and progenitor cells (data were derived from GSE42414, Laurenti et al, Nature Immunology, 2013).


A
% of CD45+GFP+ in blood
EV
CBX7
6 weeks
10 weeks
14 weeks
18 weeks
*
p=0.015
B
% of huCD45+GFP+ in BM
*
p=0.034
28-40 weeks
C
% of CD45+GFP+ in blood
*
p=0.026
D
% of huCD45+GFP+ in BM
**
p= 0.008
18,22,33 weeks
E
CD3
CD19
Others
CD33
F
% of CD33+ cells in CD45+GFP+ in BM
**
p= 0.003
CD33
CD45
59.2 %
94.0 %
G
% of CD38- in CD34+LIN-CD45+GFP+ in BM
*
p= 0.035
CD38
CD34
5.38 %
8.31 %

mediated knockdown of *CBX7* in CD34$^+$ HSPCs with two distinct short hairpins. Indeed, knockdown of *CBX7* resulted in a 3-fold reduced LTC-IC frequency (Supplementary Figure 1B).

These *in vitro* phenotypes prompted us to analyze endogenous *CBX* expression levels in different primitive hematopoietic cell subsets using previously published microarray experiments (Laurenti et al., 2013). *CBX7* expression decreased during differentiation from HSCs (HSC1 = Lin$^-$ CD34$^+$CD38$^-$CD45RA$^-$CD90$^+$CD49f$^+$, HSC2 = Lin$^-$ CD34$^+$CD38$^-$ CD45RA$^-$CD90$^-$CD49f$^+$) to more mature MPPs, CMPs and GMPs subsets (Supplementary Figure 1C).

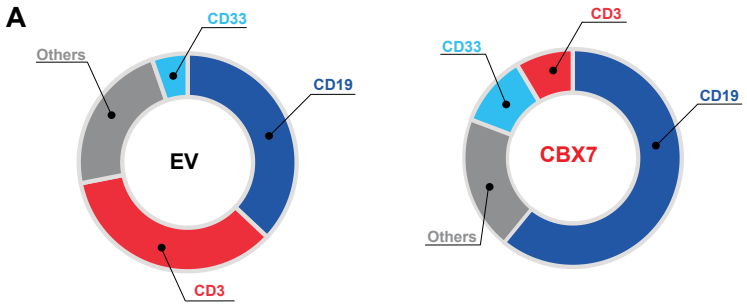
CBX7 overexpression enhances engraftment of human CD34$^+$ cord blood *in vivo*

As CBX7 proved to be the most potent inducer of CD34$^+$ HSPC proliferation *in vitro*, we assessed whether its overexpression improved engraftment of these cells *in vivo.* To this end, we transduced CD34$^+$ cord blood cells with a *CBX7* expressing vector and transplanted the equivalent of 2 × 10^5 CD34$^+$GFP$^+$ cells in sub-lethally irradiated female NOD-SCID

Figure 2:

***CBX7* overexpression induces enhanced long-term engraftment, myelopoiesis and self-renewal of primitive CD34$^+$CD38$^-$ HSPCs *in vivo*.**

(**A**) Human chimerism levels in the peripheral blood of NSG mice upon transplantation of 200,000 CD34$^+$GFP$^+$ *CBX7* overexpressing or *EV* control cord blood cells. (**B**) Primary recipients were sacrificed after 28-40 weeks, and human engraftment in bone marrow was evaluated. (**C**) Human chimerism levels in the peripheral blood of NSG mice upon transplantation of 7 days ex vivo cultured *CBX7* overexpressing or empty vector control CD34$^+$ cord blood cells. (**D**) Human engraftment in bone marrow of mice shown in panel C, combined analysis of mice sacrificed after 18, 22 and 33 weeks post-transplant. (**E**) Relative engraftment of human CD19$^+$, CD3$^+$ and CD33$^+$ cells within human CD45$^+$GFP$^+$ bone marrow cells of mice shown in panel C, combined analysis of mice sacrificed after 18, 22 and 33 weeks post transplantation. (**F**) Human myeloid engraftment in bone marrow of mice shown in panel C, combined analysis of mice sacrificed after 18, 22 and 33 weeks post transplantation. (**G**) Frequency of CD34$^+$CD38$^-$ cells in huCD45$^+$GFP$^+$lin$^-$CD34$^+$ cells in bone marrow of mice shown in panel C, combined analysis of mice sacrificed after 18, 22 and 33 weeks post transplantation. (Identically colored circles indicate paired experimental and control samples that originate from the same cord. Statistical analysis was performed using a two-tailed Mann-Whitney test).



Supplementary Figure 2, related to Figure 2

(**A**) Relative engraftment of human CD19+, CD3+ and CD33+ cells within the human $CD45^+GFP^+$ peripheral blood cell fraction, measured 18 weeks post transplantation of *CBX7* or *EV* overexpressing cells into NSG mice (EV n=10, CBX7 n=14, two cords).

IL2rγnull (NSG) mice. After transduction, cells were kept in culture for 24 hours before transplantation. We used three freshly isolated cords and from 6 weeks post-transplantation onwards, we measured chimerism every 4 weeks in the peripheral blood (Figure 2A). Mice transplanted with *CBX7* overexpressing $CD34^+$ cord blood cells showed significantly higher engraftment of $CD45^+GFP^+$ cells in peripheral blood 18 weeks after transplantation. After 28-40 weeks mice were sacrificed and engraftment of GFP^+ cells in the bone marrow was analyzed. Mice transplanted with *CBX7* overexpressing cord blood cells showed ~3 fold (EV 10%, CBX7 29.5%) higher bone marrow engraftment compared to mice transplanted with the EV-control (Figure 2B).

To explore whether *CBX7* overexpression would be able to maintain human $CD34^+$ HSPCs in a more primitive state for a longer period *ex vivo*, we prolonged total *in vitro* culture time from 3 to 7 days and transplanted the equivalent of 1.5×10^6 GFP^+CD34^+ cord blood cells in irradiated NSG mice. Mice transplanted with *CBX7* overexpressing cord blood cells displayed significantly higher engraftment in peripheral blood after 18 weeks (Figure 2C). Mice were sacrificed after 18, 22, or 33 weeks and bone marrow cells were analyzed for the presence of human donor-derived cells. Overall, mice transplanted with *CBX7* overexpressing cells showed significantly higher levels of GFP^+ cells in the bone marrow (Figure 2D). Furthermore, mice transplanted with *CBX7* overexpressing $CD34^+$ cells showed a reduced percentage of $CD3^+$ T-cells and a significantly increased percentage of $CD33^+$ cells in the human $CD45^+GFP^+$

compartment in the bone marrow suggesting that overexpression of *CBX7* enhances myelopoiesis roughly twofold (Figure 2E and F). Similar results were obtained in the peripheral blood after 18 weeks (Supplementary Figure 2A). Furthermore, these mice showed a significantly higher percentage of primitive $CD38^-$ cells in the GFP^+ Lin^-CD34^+ compartment (Figure 2G), indicating that CBX7 controls *in vivo* proliferation or maintenance of human primitive hematopoietic stem- and progenitor cells.

Genome-wide transcriptional consequences of *CBX7* and *-8* overexpression

We next assessed the effect of *CBX7* and *CBX8* overexpression on the transcriptional program of human $CD34^+$ HSPCs. We used $CD34^+$ cord blood cells from 5 female newborns and transduced these with *CBX7*, *CBX8* or an empty vector control, sorted 100,000 $CD34^+GFP^+$ cells 96 hours post-transduction and performed transcriptome analysis using high-throughput RNA-sequencing.

Differential expression analysis showed a total of 1463 genes significantly up- and 1183 genes significantly down-regulated when these five independent replicate cords with *CBX7* overexpression were compared with controls.

To annotate *CBX7*-induced up- and down-regulated genes we first used traditional GO enrichment analysis, which reports significantly enriched categories of genes grouped by molecular function or biological process. Screening downregulated genes after *CBX7* overexpression for enriched biological processes revealed that more than 100 genes were associated with "cell differentiation", including "leucocyte differentiation", "lymphocyte differentiation", "epithelial cell differentiation", "T and B cells differentiation", "myeloid cell differentiation", "neuron differentiation", and "macrophage differentiation" (Figure 3A). Furthermore, we found repression of genes associated with cell cycle arrest and negative regulation of cell cycle. In contrast, *CBX7*-induced upregulated genes revealed transcripts related to cell cycle ("cell cycle", "cell cycle process", "G1/S transition of mitotic cell cycle") (Supplementary Figure 3A) and DNA replication ("DNA replication initiation", "DNA replication", "DNA conformation change", "G1/S transition of mitotic cell cycle", "chromatin assembly or disassembly") (Supplementary Figure 3B). The list of upregulated genes did not contain any GO-group associated with differentiation

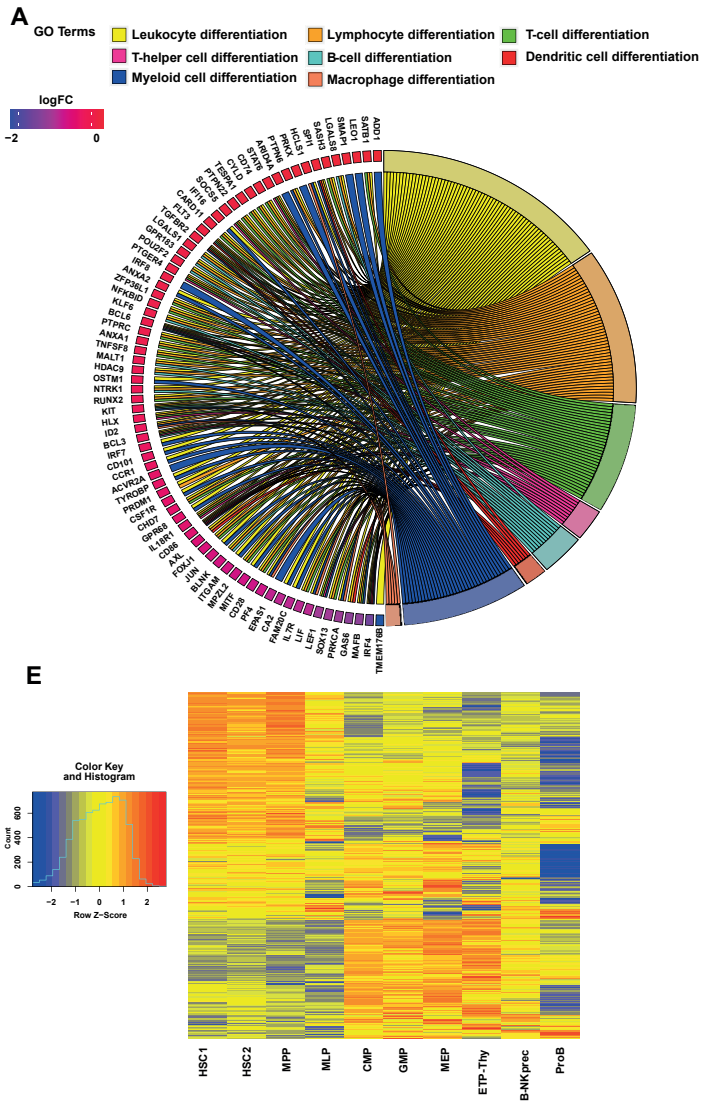
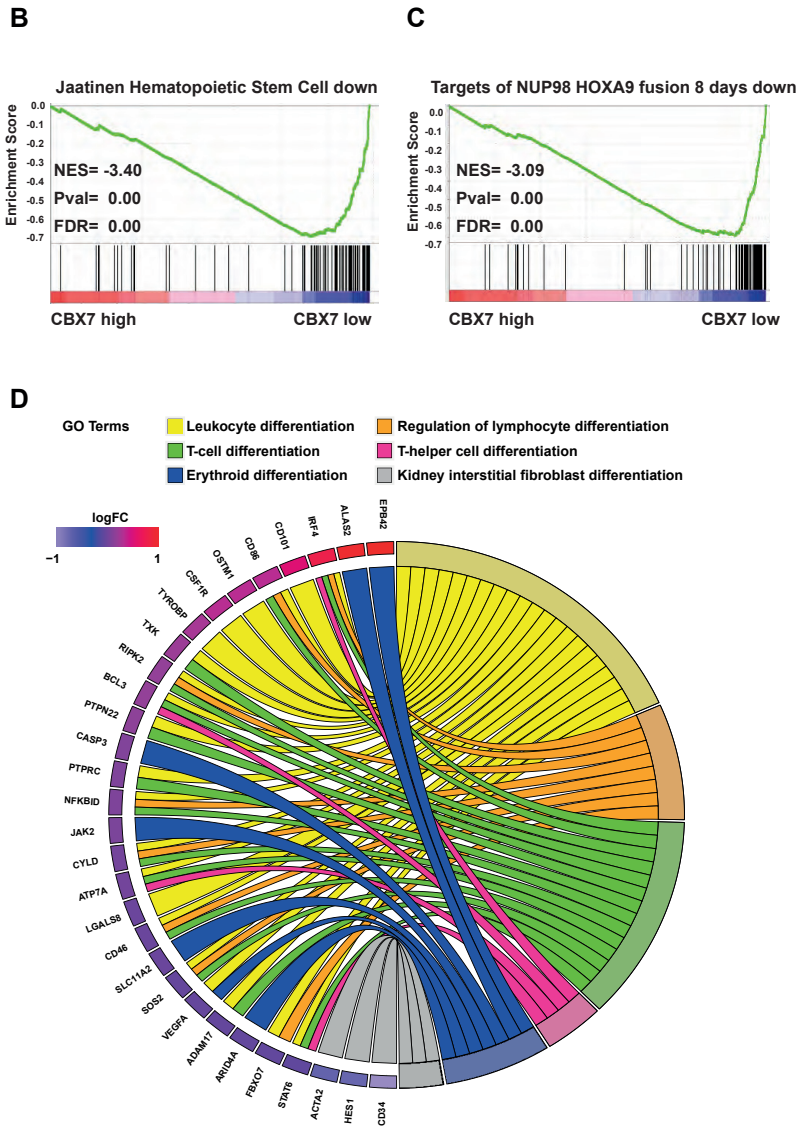


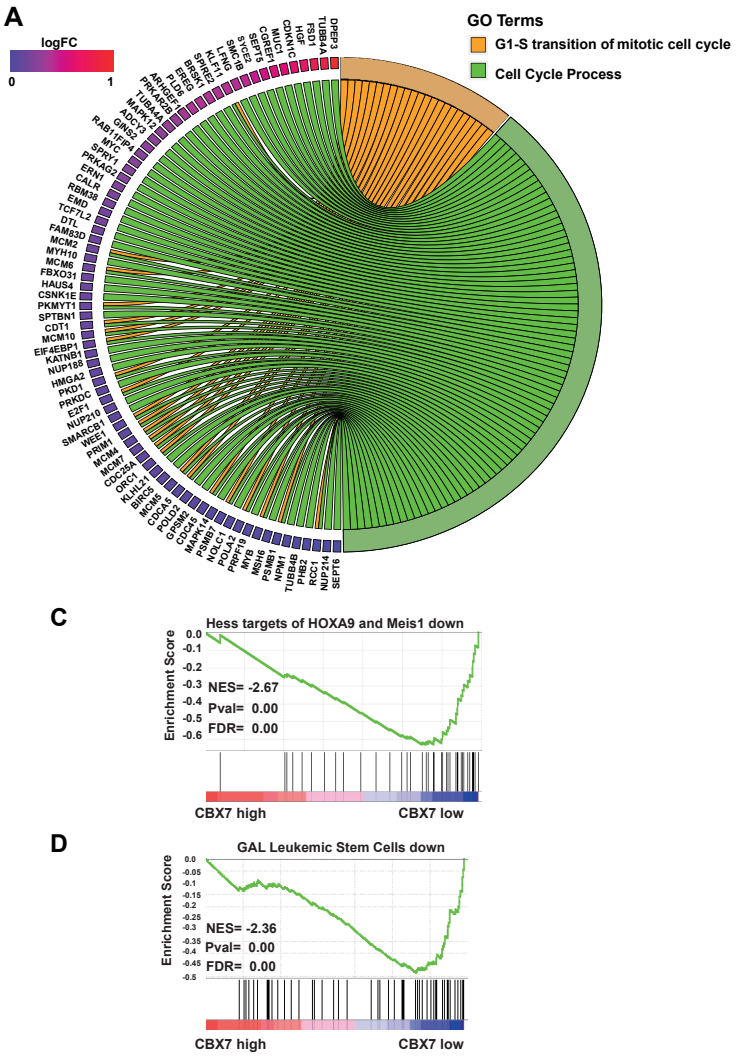
Figure 3:

RNA-Seq analysis of *CBX7*, *CBX8* and *EV* overexpressing CD34+ HSPC.

Sets of differentially expressed genes were screened for Gene Ontology (GO) enrichment. GO categories were enriched for "differentiation", "cell cycle", "chromatin" and "DNA", shown using GO Chord plots. Preranked gene set enrichment analysis was performed for differentially expressed genes (FDR<0.1) upon overexpression of *CBX7* in comparison to empty vector control cells. (**A**) GO Chord plot of genes repressed upon overexpression of *CBX7* in comparison to control cells, associated with



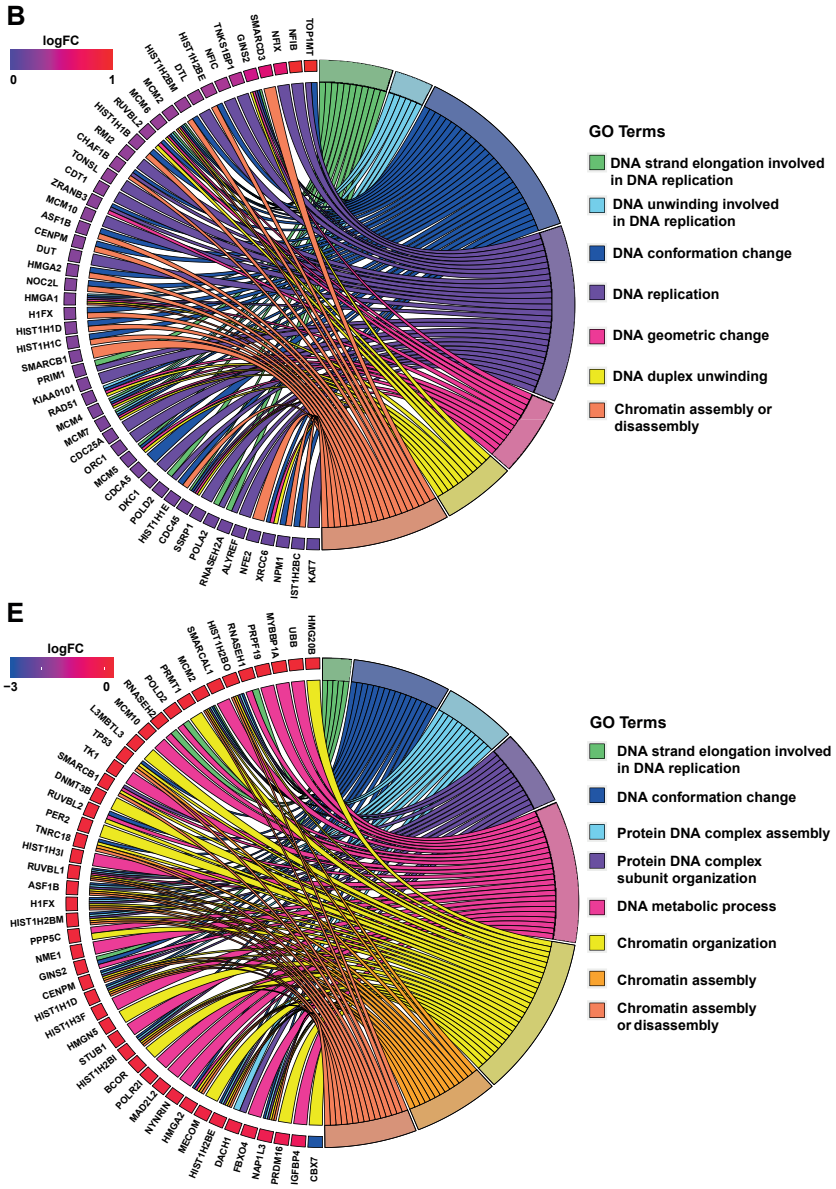
the GO terms "differentiation" of various hematopoietic cells. (**B** and **C**) Gene Set Enrichment plots for 2 out the top 3 gene sets (p<0.001) with the highest enrichment in genes downregulated upon overexpression of *CBX7* compared to control values. (**D**) GO Chord plot of genes differentially expressed upon overexpression of *CBX7* in comparison to *CBX8* overexpressing CD34+ HSPCs associated with GO terms "differentiation" of various cell types. (**E**) Heatmap containing genes upregulated upon overexpression of *CBX7* and their expression in multiple normal hematopoietic subsets according to previously published data from (Laurenti et al. Nature Immunology, 2013).



Supplementary Figure 3, related to Figure 3

Sets of differentially expressed genes were screened for Gene Ontology (GO) enrichment.

(**A**) GO categories for genes upregulated upon overexpression of *CBX7* in comparison to empty vector cells were enriched for "G1-S transition of mitotic cell cycle" or "cell cycle process". GO Chord plot is depicted of genes upregulated upon overexpression of *CBX7* in comparison to control cells, associated with the GO terms "G1-S transition of mitotic cell cycle" or "cell cycle process". (**B**) GO categories for genes upregulated upon overexpression of *CBX7* in comparison to *EV* overexpressing CD34+ HSPCs were enriched for GO-terms containing "DNA" or "chromatin". GO Chord plot is depicted of genes up-regulated upon overexpression of *CBX7* in comparison to *EV* overexpressing CD34+ HSPCs, associated



with the GO terms "DNA" or "chromatin".(**C, D**) Preranked Gene Set Enrichment analysis was performed for differentially expressed genes (FDR<0.1) upon overexpression of *CBX7* in comparison to empty vector control cells. Enrichment-plots of two gene-sets (p<0.001) are shown that are enriched in genes downregulated upon overexpression of *CBX7* in comparison to EV. (**E**) GO Chord plot illustrating genes differentially expressed upon overexpression of *CBX7* in comparison to *CBX8* overexpressing CD34+ HSPCs, associated with GO-terms containing "DNA" and "chromatin".

of any kind of hematopoietic cells. These GO annotations are in good agreement with the *in vitro* and *in vivo* observation that overexpression of *CBX7* leads to elevated self-renewal.

The transcriptional consequences of *CBX7* overexpression were —as anticipated— complex, but revealed perturbations of multiple genes known to be crucial for HSC behavior. In total, there were 146 genes downregulated upon *CBX7* overexpression which were related to "transcription". From those, 36 were transcription factors, and 22 genes were related to histone modifications, 2 of which were Polycomb associated genes (*BMI1* and *SCML1*).

Reversely, genes upregulated upon *CBX7* overexpression revealed 160 genes related to "transcription". From those, 19 genes contained the term "transcription factor" (including *GATA1*, *CITED2*, *RUNX1*, *MYC*, *HOXA7*). No Polycomb genes were upregulated (except for *CBX7*). Another 19 genes were related to the term "histone modifications" (including *DNMT3A* and *KDM1A*). Further, we observed upregulation of genes important for myelopoiesis, including *CEBPA, MPO* and the *G-CSF-receptor*.

Complementary, we performed Gene Set Enrichment Analysis (GSEA) on a pre-ranked list containing all genes differentially expressed (FDR<0.1) upon *CBX7* overexpression in comparison to the empty vector control. We sorted gene-sets with a FDR<0.25 and p<0.01 according to their normalized enrichment-score (NES). Interestingly, GSEA revealed a strong negative correlation (high NES) with a gene set containing genes with low abundance in $CD133^+$ HSCs, indicating that increased levels of *CBX7* results in maintained repression of genes which are usually barely expressed in HSCs (Figure 3B). Furthermore, we identified two other sets with a high negative correlation, both containing genes downregulated upon overexpression of *HOXA9* either with *NUP98* or *Meis1*, suggesting that CBX7 targets overlap with targets of these fusion oncogenes (Figure 3C and Supplementary Figure 3C). Furthermore, we found a strong negative correlation with a gene set containing genes lower expressed in leukemic stem cells ($CD34^+CD38^-$) in comparison to leukemic blasts ($CD34^+CD38^+$) suggesting that genes downregulated by *CBX7* overexpression are indeed lower expressed in immature leukemic stem cells than in more differentiated leukemic blasts (Supplementary Figure 3D).
Transcriptome analysis of *CBX8* overexpressing $CD34^+$ cells resulted in 1444 significantly upregulated and 815 downregulated genes in comparison to empty vector. As the cell biological consequences of *CBX7* and

CBX8 overexpression were quite distinct, we compared differential gene expression patterns of *CBX7* with *CBX8* and identified a fraction of genes specifically up- (334) or down-regulated (346) by either *CBX7* or *CBX8*. Interestingly, GO-analysis of differentially expressed genes upon *CBX8* overexpression in comparison to EV revealed similar suppression of differentiation pathways. However, lymphoid pathways were suppressed to a lesser extent, and many of the replication/cell cycle genes were missing from the list of upregulated genes (Supplementary Figure 3E). GO-Analysis for genes differentially expressed between *CBX7* and *CBX8* overexpression conditions revealed that genes specifically upregulated upon overexpression of *CBX8* in comparison to *CBX7* were associated with differentiation of hematopoietic cells (Figure 3D, Supplementary Figure 3B). These genome-wide transcriptome analyses are in accordance with *CBX7* overexpressing $CD34^+$ cells having a higher CFU-replating efficiency and a higher CAFC-frequency compared to *CBX8*.

To further characterize differentially expressed genes upon *CBX7* overexpression, we compared these with steady state transcriptomes of multiple subsets of hematopoietic cell types, using a previously published expression data set as a cross reference (Laurenti et al., 2013). This analysis revealed that 378 transcripts that were higher expressed upon *CBX7* overexpression were preferentially abundant in the more primitive cell compartments (HSC1, HSC2, MPP versus MLP, CMP, GMP, MEP, ETP-Thy, B-NKprec, ProB) (Figure 3E). This suggests their involvement in maintaining elevated levels of self-renewal upon overexpression of *CBX7* in HSCs.

This group of primitive-signature genes includes *CCND2*, *ERG*, *FLI1*, *HMGA2*, *IGF1R*, *LMO2*, *MEIS1*, *MYCN*, *PER2*, *PTK2*, *RUNX1*, *SPINT1*, *ZBTB16*, and *ZEB1*, and all belong to the KEGG pathway GO group "Transcriptional misregulation in cancer". Also, overexpression of *CBX7* resulted in downregulation of *CD38* and upregulation of *CD34*, two markers which are used for identifying primitive hematopoietic cells in FACS stainings.

In summary, our transcriptome analysis clearly reveals that CBX7 mediates its activity through repression of genes important for differentiation of hematopoietic cells and upregulation of genes important for cell cycle. These include multiple well-known upregulated oncogenes and downregulated tumor suppressor genes.

CBX7 expression is elevated in AML and its repression result in differentiation of AML cells.

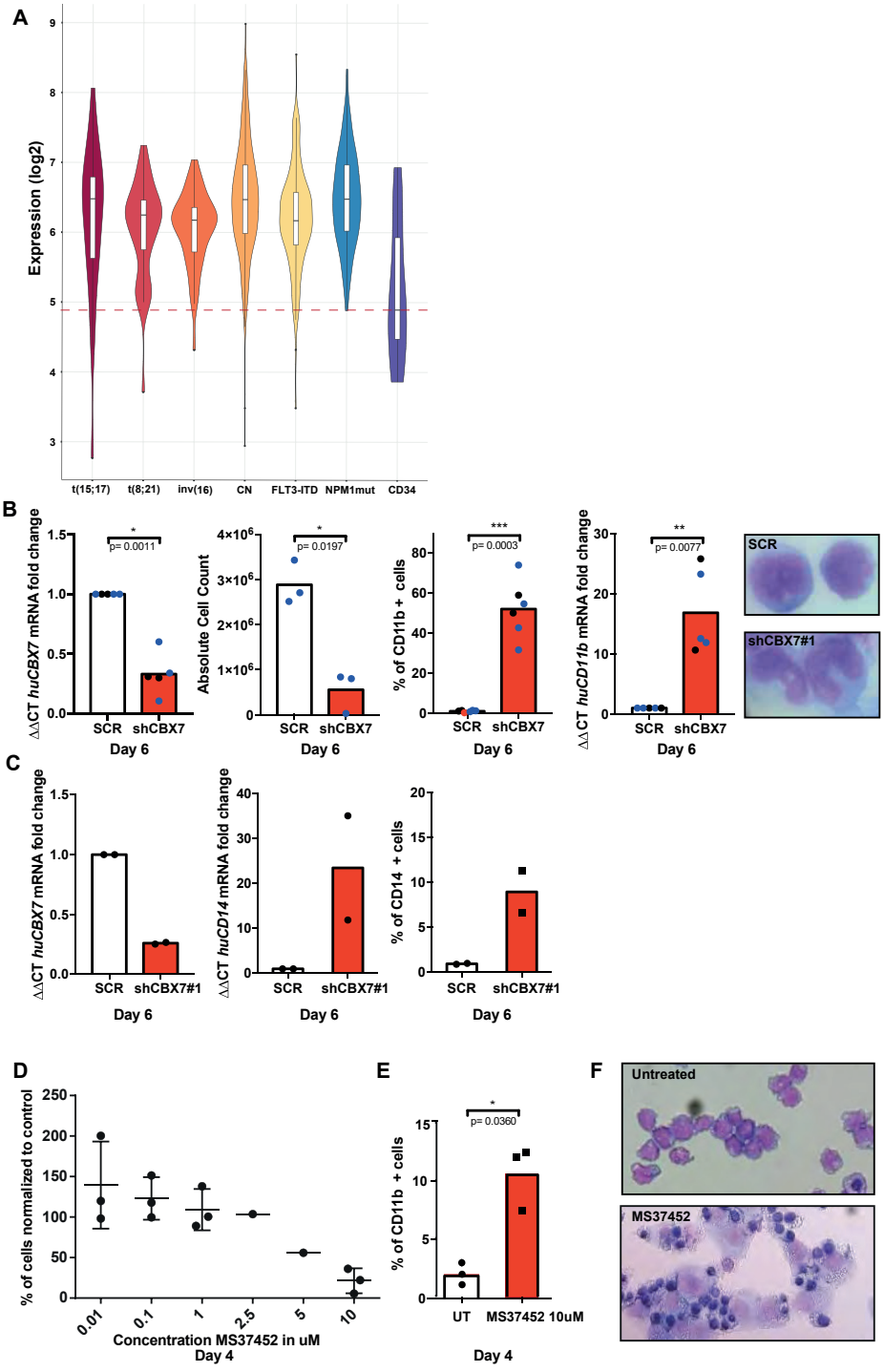
Our data show that CBX7 is able to increase self-renewal of normal human hematopoietic stem and progenitor cells. To explore a putative role for CBX7 in the maintenance of AML cells, we first analyzed *CBX7* mRNA expression levels in AML patient samples in two previously published data sets. In the first data set, containing 529 AML patient samples from patients treated at the Erasmus MC (Rotterdam, The Netherlands), *CBX7* expression was significantly upregulated in comparison to peripheral blood mobilized $CD34^+$ cells (Verhaak et al., 2009). The highest expression was observed in acute promyelocytic leukemia (t(15;17); APL), leukemias with a normal karyotype and *NPM1* mutated leukemia (Figure 4A). We additionally analyzed data from the Cancer Genome Atlas via Bloodspot (Bagger et al., 2016). Also in this patient cohort, *CBX7* level was significantly higher in multiple AML subtypes (Supplementary Figure 4A).

To explore a functional role for high *CBX7* expression in human leukemia, we assessed to what extent depletion of *CBX7* would affect leukemic cell growth. As *CBX7* is more abundantly expressed in APL (Figure 4A),

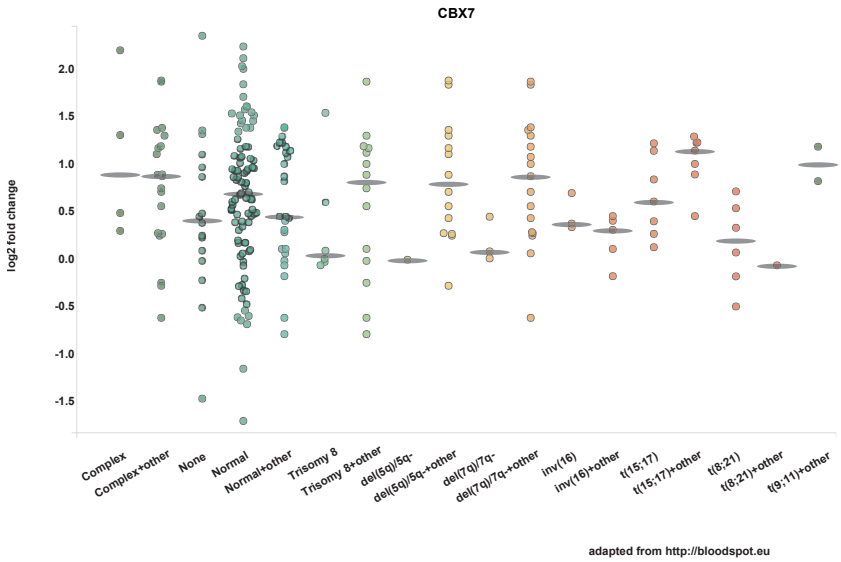
Figure 4:

CBX7 is significantly higher expressed in AML patient samples, and its knockdown induces differentiation in AML cells.

(**A**) Analysis of CBX7 expression in 529 AML patient samples by microarray. Figure 4A shows a violin plot displaying expression of various AML subtypes and CD34+ peripheral blood mobilized stem cells. (**B**) Short-hairpin mediated knockdown of *CBX7* mRNA (3B, 1st panel) in HL60 cells results in upregulation of CD11b on mRNA and protein levels (3B, 3rd + 4th panel) after six days. Multiple cells showed signs of differentiation upon knockdown of *CBX7*. (3B, 5th panel) (black= shCBX7#1, blue = shCBX7#2). (**C**) Short-hairpin mediated knockdown of CBX7 mRNA (3C, 1st panel) in OCI-AML3 cells results in upregulation of CD14 on mRNA (3C, 2nd panel) and protein levels (3C, 3rd panel) after six days. (**D**) Growth of OCI-AML3 cells treated with the CBX7 chromodomain inhibitor MS37452 in different concentrations after four days in culture. (**E**) Treatment of OCI-AML3 cells with MS37452 at a concentration of 10μM results in increased expression of CD11b. (**F**) MS37452 induces monocyte/macrophage differentiation in OCI-AML3 cells. After treatment for 4 days with MS37452 at concentration of 10μM, cytospin preparations were stained with May-Grünwald Giemsa stain. Magnification 40x


A
Expression (log2)
t(15;17)
t(8;21)
inv(16)
CN
FLT3-ITD
NPM1mut
CD34
B
ΔΔCT huCBX7 mRNA fold change
p= 0.0011
SCR
shCBX7
Day 6
Absolute Cell Count
p= 0.0197
% of CD11b + cells
p= 0.0003
ΔΔCT huCD11b mRNA fold change
p= 0.0077
SCR
shCBX7#1
C
ΔΔCT huCBX7 mRNA fold change
ΔΔCT huCD14 mRNA fold change
% of CD14 + cells
D
% of cells normalized to control
Concentration MS37452 in uM
Day 4
E
% of CD11b + cells
p= 0.0360
UT
MS37452 10uM
Day 4
F
Untreated
MS37452

A



Supplementary Figure 4, related to Figure 4

(**A**) Log2 fold change of *CBX7* mRNA expression in the AML TCGA dataset analyzed with the nearest normal counterpart method (Rapin et al, Blood 2014). (Source http://bloodspot.eu)

we downregulated *CBX7* mRNA using a short-hairpin approach in HL60 cells, which harbor a t(15;17) translocation. Knockdown of *CBX7* was associated with a reduced abundance of *CBX7* mRNA to ~40% of normal levels (Figure 4B, first panel) and lower absolute cell numbers after 6 days in culture (Figure 4B, second panel). Strikingly, downregulation of *CBX7* resulted in a significant increase of CD11b expression, which is usually not expressed on primitive APL-blasts but rather on mature monocytes, macrophages and granulocytes (Figure 4B, third and fourth panel). The changes of CD11b protein levels were associated with an increased expression of *CD11b* on mRNA level (Figure 4B, fourth panel), and morphological signs of cellular maturation upon May-Grünwald Giemsa staining (Figure 4 B, fifth panel).

It has been reported that CBX7 can interact with mutated DNMT3A(R882), but not with wild type DNMT3A in AML patient samples (Koya et al., 2016). Therefore, we decided to downregulate *CBX7* in OCI-AML3 cells, a cell line carrying *DNMT3A* R882 and mutant *NPM1*.

Similar as in HL60 cells, upon knockdown of *CBX7*, OCI-AML3 cells started to differentiate and upregulated the differentiation marker CD14 on protein and mRNA level (Figure 4C). In summary, these experiments indicate that CBX7 is necessary for maintaining leukemic cells in an undifferentiated state, independent of DNMT3A.

We tested whether pharmacological inhibition of CBX7 would result in similar effects as short hairpin mediated repression. To this end we cultured OCI-AML3 cells in the presence of increasing concentrations of the small molecule MS37452, which has been shown to bind to residues in the chromodomain of CBX7 so that protein-protein interactions are disturbed. This loss of normal chromodomain function resulted in derepression of PRC target genes in prostate cancer cells (Ren et al., 2015). In OCI-AML3 cells MS37452 resulted in loss of cell growth in a time- and dose-dependent manner. Furthermore, MS37452 treatment induced differentiation in leukemic cells, as evidenced by upregulation of the differentiation marker (and CBX7 target) CD14 and by the strong increase of cells with a highly differentiated morphology (Figure 4E and F).

CBX7 interacts with trimethylated non-Polycomb proteins

To further unravel the molecular mechanism by which CBX7 exerts its potent activity and taking into account that PcG proteins are known to operate in large protein complexes, we decided to identify proteins directly interacting with CBX.

We performed label-free mass-spectrometry analysis of benzoase treated proteins that co-precipitated with FLAG-tagged CBX7, FLAG-tagged CBX8 and FLAG-tagged CBX4, using murine and human cells. A protein fraction that co-precipitated with FLAG-tagged GFP was used as a negative control. To prioritize candidates, we first removed proteins with low spectral counts (<10% of the cumulative spectral count) and then ranked proteins in relation to their spectral counts. We compared all MS sets and screened for consistent binding partners of both murine and human CBX proteins. As expected, multiple members of PRC1 and -2 complexes were identified, including PCGF1, PCGF2, PCGF6, SCML2, PHC1, PHC2, PHC3, BMI1, RING2, RING1, EED, PCGF5, and SUZ12. The finding of those known CBX interaction partners confirmed that other proteins in the pull downs could be potential members of CBX-containing protein complexes. Interestingly, a considerable number of histone modifiers (63 proteins), transcription

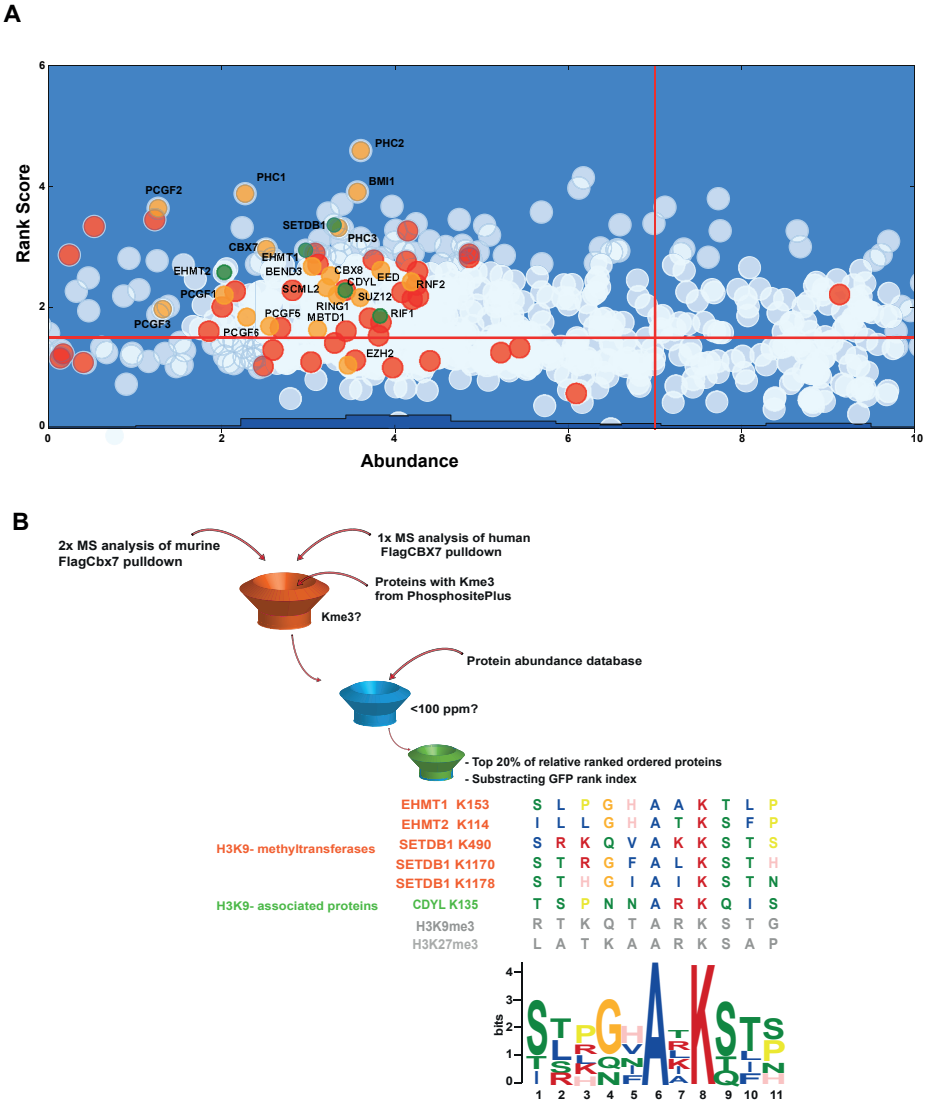
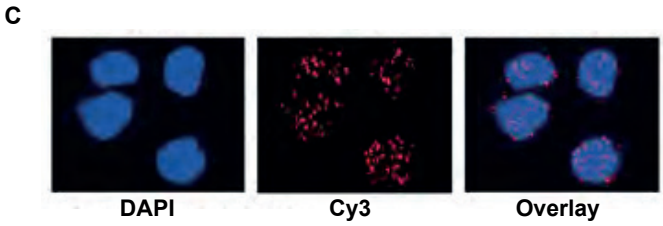
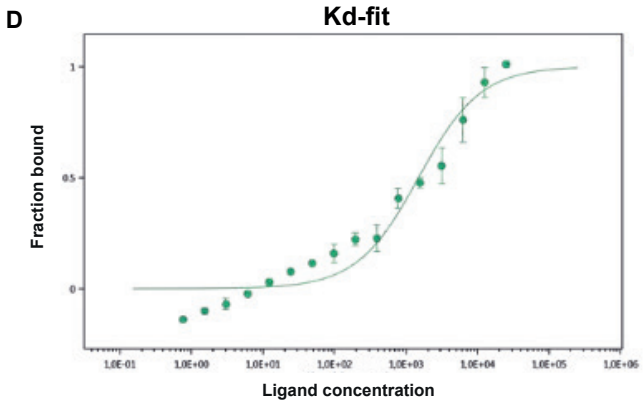


Figure 5:

Mass spectrometry analysis of FLAG-pulldowns reveals multiple H3K9 methyltransferases as CBX7-binding partners harboring a trimethylated-lysine embedded in a motif highly similar to H3K9me3 and H3K27me3

(**A**) Search for putative interaction partners of CBX proteins by label-free mass spectrometry. The 2D plot depicts CBX7 interaction partners ranked on the basis of their cumulative rank score (derived from the frequency of spectral counts, corrected for GFP control samples) and the average abundance of these proteins in the human PaxDB database. The top-left corner represents priority candidates.

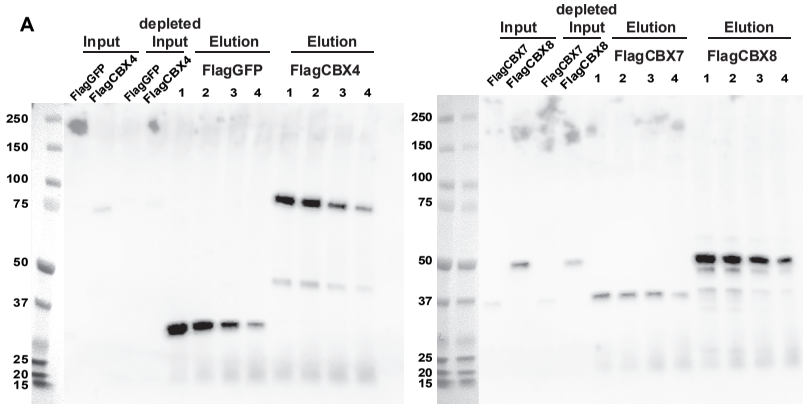


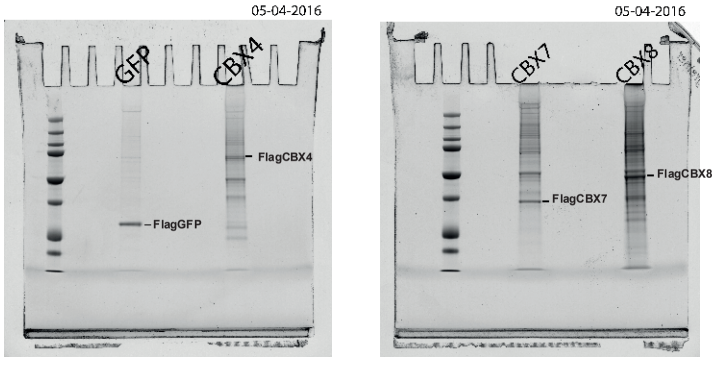


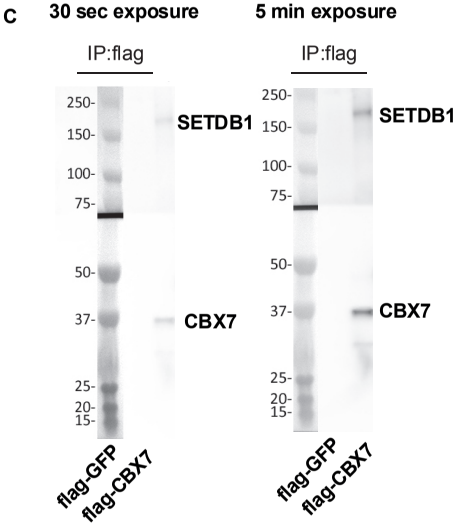
Red symbols indicate known transcription factors, orange symbols refer to Polycomb proteins, and green symbols are proteins with known lysine trimethylation sites. (**B**) Schematic overview of the analysis resulting in the identification of tri-methylated CBX7-binding proteins, the corresponding peptides and their CBX7 tri-methylated binding motif. (**C**) Duolink proximity ligation assay (PLA) of endogenous SETDB1 and CBX7 performed in HL60 cells. Each PLA signal (Cy3, red) is indicative of one detected interaction event. Nuclei (blue) are stained with DAPI. (**D**) Dissociation constant (Kd)-fit of labelled SETDB1 used at a constant concentration of 25nM and (unlabelled) CBX7 titrated until a final concentration of 25μM (1:1) determined by Microscale Thermophoresis (MST).

factors (20 proteins: GTF3C1, YY1, ZFPM1, MGA, GATA1, BPTF, GTF3C5, AHCTF1, LDB1, ADNP, ELF2, GTF3C3, PELP1, E4F1, CBFA2T3, GTF3C4, PRDM2, NCOR1, CDYL, HLTF) and DNA repair associated proteins (DDB1, RPS3, RFC1, MSH6, LIG3, TRIP12, BRCA2, RNF169, MRE11A, INO80, BRCA1, POLB, EPC2) were also detected (Figure 5 A).

Since canonically CBX7 binds to the trimethylated lysine of H3K27 through its chromodomain, we hypothesized that the chromodomain could potentially associate with other trimethylated lysines in non-histone proteins when they contain a peptide context similar to H3K27.


A
depleted
Input
Input
Elution
Elution
FlagGFP
FlagCBX4
FlagGFP
FlagCBX4
1 2 3 4
1 2 3 4
FlagGFP
FlagCBX4
250
150
100
75
50
37
25
20
15
depleted
Input
Input
Elution
Elution
FlagCBX7
FlagCBX8
FlagCBX7
FlagCBX8
1 2 3 4
1 2 3 4
FlagCBX7
FlagCBX8
250
150
100
75
50
37
25
20
15


B
05-04-2016
GFP
CBX4
FlagCBX4
FlagGFP
05-04-2016
CBX7
CBX8
FlagCBX8
FlagCBX7


C
30 sec exposure
5 min exposure
IP:flag
IP:flag
250-
150-
100-
75-
50-
37-
25-
20-
15-
SETDB1
CBX7
flag-GFP
flag-CBX7
250-
150-
100-
75-
50-
37-
25-
20-
15-
SETDB1
CBX7
flag-GFP
flag-CBX7

Therefore, we screened the list of CBX7 human and murine binding partners for proteins harboring a putative trimethylated lysine using the PhosphoSitePlus database (Hornbeck et al., 2015). This screen revealed a list of 218 human and murine trimethylated proteins, corresponding to 335 known trimethylated human and murine peptides. We only considered proteins with high spectral counts (top 15% of the relative rank ordered proteins) and low protein abundance (<100 ppm, pax-db.org, average of all samples) and corrected for the binding of each candidate to GFP. This strict filtering narrowed our list down to four proteins (CDYL, SETDB1, EHMT1 and EHMT2) (Figure 5B). We then applied the same filtering for murine Cbx7-binding proteins and identified SETDB1, EHMT1 and EHMT2 as evolutionary conserved binding partners. When using less strict filtering rules (top 70%) we also identified CDYL as a binding partner of murine Cbx7. Interestingly, this list of trimethylated CBX7 interaction partners contains three H3K9 methyltransferases (EHMT1, EHMT2, SETDB1) and one H3K9me3 associated proteins (CDYL). Next to our approach to use spectral counts for identifying putative interaction partners, we also calculated the relative enrichment in the exponentially modified protein abundance index (emPAI) in the CBX7-sample over the control-sample (FlagGFP). All our candidate proteins (SETDB1, EHMT1, EHMT2, CDYL) showed a relative enrichment of at least 12.

We next assessed whether these binding partners to human CBX7 contain a common signature. Indeed, the consensus-binding motif for CDYL, EHMT1, EHMT2, SETDB1 was found to be S[LT]PGHA.Kme3ST[PS], which is highly similar to the peptide sequences to which human CBX7 is known to bind: A[RILFYV]Kme3[ST], (Kaustov et al., 2011) (Figure 5B). The similarity of these motifs suggests that CBX7 interacts with these non-histone proteins via the chromodomain. To validate the interaction of SETDB1 and CBX7 in FLAG-tagged CBX7 overexpressing K562 cells

Supplementary Figure 5, related to Figure 5

(**A**) Western blot of nuclear lysates of K562 cells after FLAG-immunoaffinity purification of FLAG-CBX7, FLAG-CBX8, FLAG-CBX4 and FLAG-GFP. (**B**) Coomassie-stained SDS-polyacrylamide gels of FLAG-CBX7, FLAG-CBX8, FLAG-CBX4 and FLAG-GFP purification Eluates 1-3 were combined. (**C**) Western blot analysis of combined eluates (1-3) upon FLAG-immunoaffinity purification of FLAG-tagged CBX7 and FLAG-tagged GFP. The upper part of the membrane was probed with an antibody against SETDB1, the lower part of the membrane was stained against CBX7.

we stained the immunoblots for CBX7 and SETDB1 (Supplementary Figure 5C).

Finally, we used two independent approaches to confirm that CBX7 and SETDB1 are indeed bona fide binding partners. First, we performed Proximity Ligation Assays (PLA) of endogenous SETDB1 and CBX7 in HL60 cells and found intense co-staining throughout the nucleus (Figure 5C). To assess whether CBX7 and SETDB1 are direct interaction partners we used Microscale Thermophoresis. In a cell-free environment these two proteins showed strong binding to each other, with a dissociation constant (K_d)-fit of labelled SETDB1 and (unlabelled) CBX7 of 1.5 μM, which in fact is much lower than the previously reported K_d-value of CBX7 to H3K27me3 (Figure 5D) (Kaustov et al., 2011).

In summary, our data show that human Polycomb CBX7 is able to interact with H3K9 methylation-associated proteins that harbor a trimethylated lysine. These interactions expand the role of CBX far beyond reading H3K27me3 and identify a non-canonical role for CBX7 in providing crosstalk to epigenetic pathways governing H3K9 methylation.

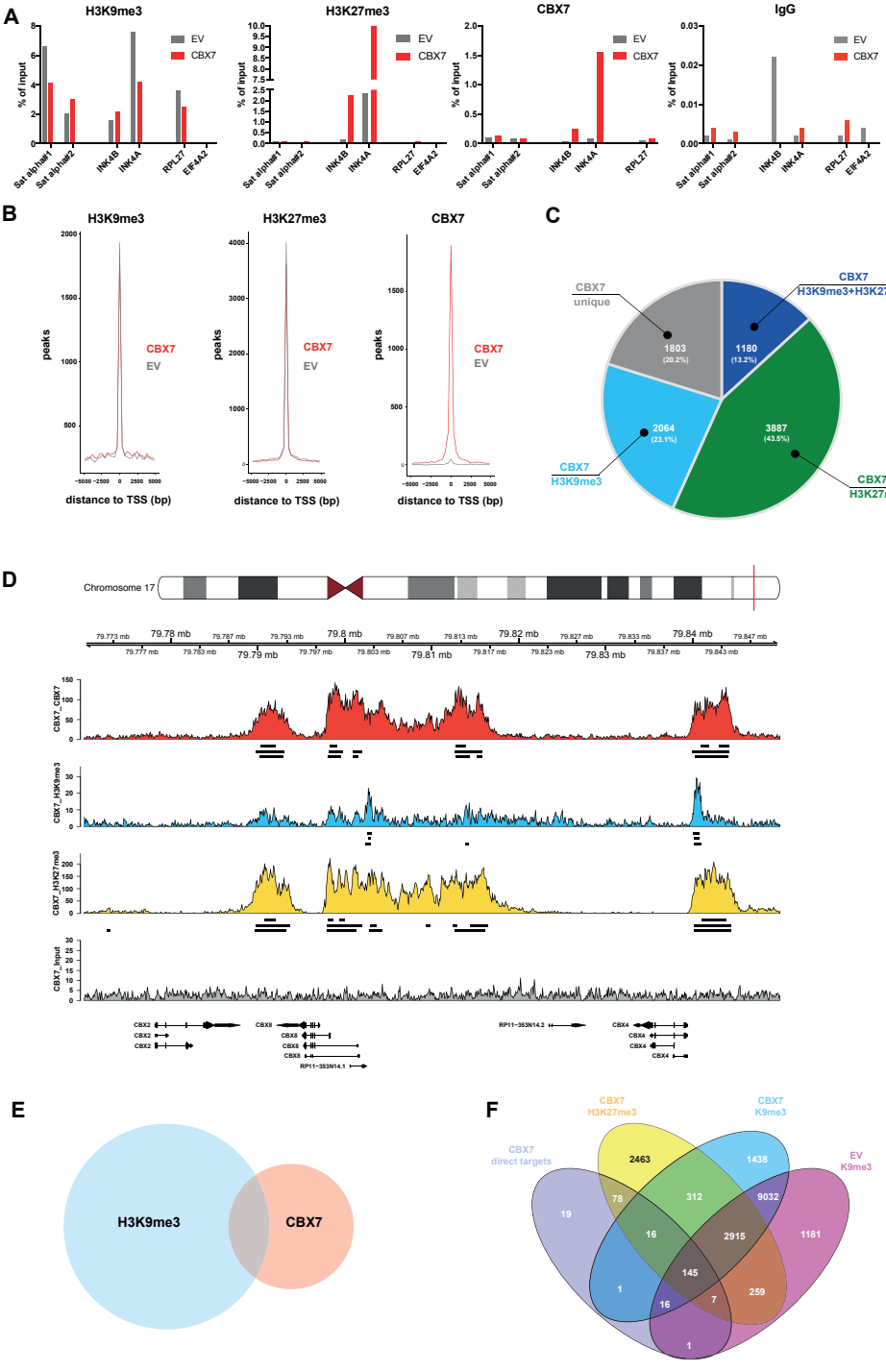
Identification of CBX7 target loci and their association with H3K9me3 and H3K27me3

To identify genes directly controlled by CBX7 and to unravel their association with repressive histone marks, we performed multiple ChIP-Seq

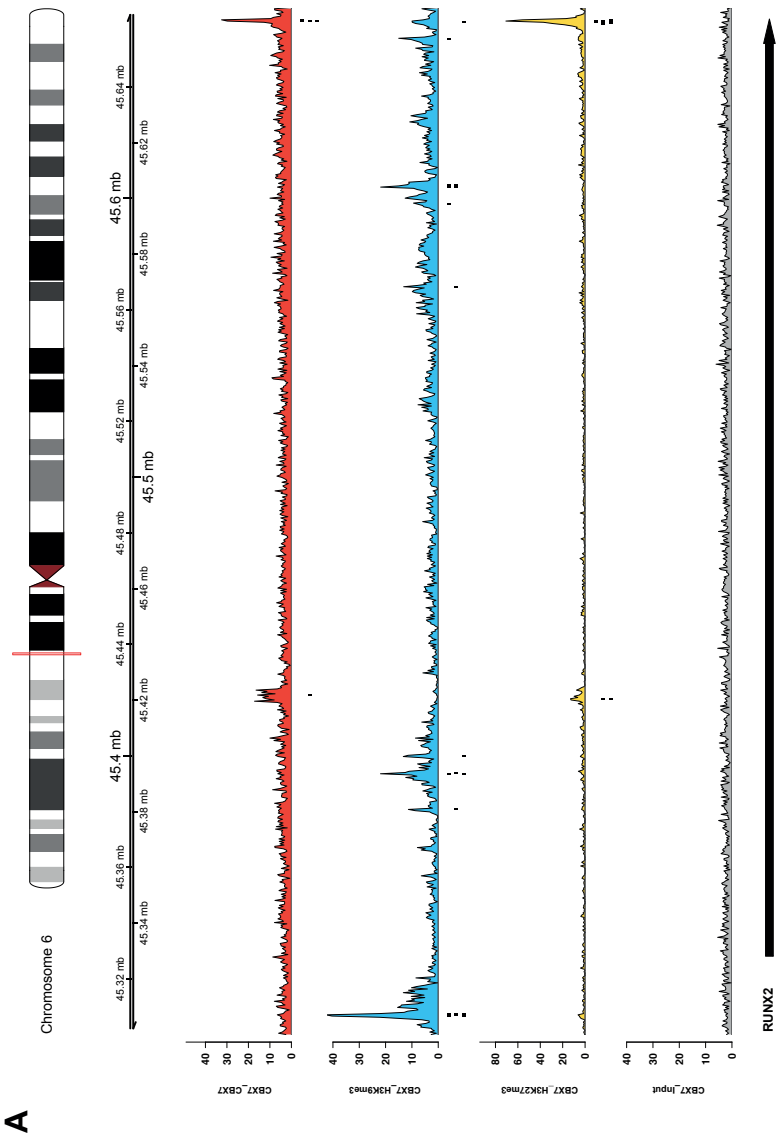
Figure 6:

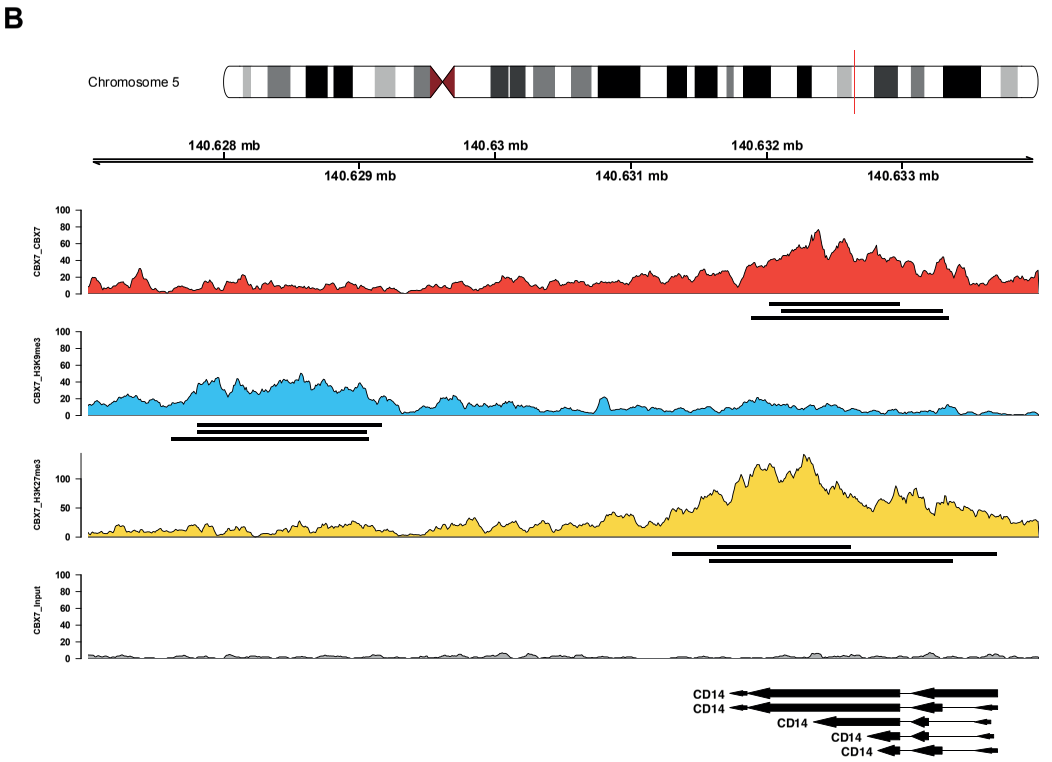
Identification of CBX7 genome-wide binding sites in primary human CD34+ cells and their association with H3K9me3 and H3K27me3.

(**A**) Chip-qPCR validation of selected positive and negative H3K9me3, H3K27me3, and CBX7 target loci, and IgG (control). (Data from one representative experiment are shown). (**B**) Genome-wide distribution of H3K9me3-, H3K27me3- and CBX7- peaks to nearest TSS in bps. (**C**) Pie-chart showing absolute and relative numbers of genome-wide CBX7 peaks and their overlap with H3K9me3 and/or H3K27me3 peaks. (**D**) Merged ChiP-Seq tracks of CBX7, H3K9me3, H3K27me3-experiments compared to input controls highlighting the genomic CBX2, CBX4 and CBX8 locus. Peak calling tracks are shown for each biological replicate individually. (**E**) Euler diagram showing overlap of TSS marked with CBX7 and H3K9me3, using a +/- 5000 bp threshold. (**F**) Venn diagram showing overlap of genes marked with H3K27me3 (CBX7 H3K27me3) and H3K9me3 (CBX7 H3K9me3) in *CBX7* overexpressing CD34+ HSPCs, H3K9me3 in control CD34+ HSPCs (EV H3K9me3) and direct targets of CBX.


A
H3K9me3
H3K27me3
CBX7
IgG
% of input
EV
CBX7
Sat alpha#1
Sat alpha#2
INK4B
INK4A
RPL27
EEF4A2
B
H3K9me3
H3K27me3
CBX7
peaks
CBX7
EV
distance to TSS (bp)
C
CBX7
unique
1803
(20.2%)
CBX7
H3K9me3+H3K27m
1180
(13.2%)
2064
(23.1%)
CBX7
H3K9me3
3887
(43.5%)
CBX7
H3K27me
D
Chromosome 17
79.78 mb
79.79 mb
79.8 mb
79.81 mb
79.82 mb
79.83 mb
79.84 mb
CBX7_CBX7
CBX7_H3K9me3
CBX7_H3K27me3
CBX7_Input
E
H3K9me3
CBX7
F
CBX7
direct targets
CBX7
H3K27me3
CBX7
K9me3
EV
K9me3
2463
1438
78
312
9032
19
16
2915
1181
145
1
16
7
259
1

experiments, all using primary transduced and expanded CD34$^+$ cord blood cells. As we found that CBX7 interacted with multiple H3K9 methyltransferases, we did not only analyze canonical H3K27me3 peaks, but also searched for loci covered with H3K9me3 upon overexpression of *CBX7* in primary cells. We pooled CD34$^+$ cord blood cells from different donors, transduced these with *CBX7* overexpressing or empty





Supplementary Figure 6, related to Figure 6
Merged ChIP-seq tracks of CBX7, H3K9me3, H3K27me3-experiments compared to input. The genomic *RUNX2*- (**A**) and *CD14* (**B**) loci are highlighted. Peak calling tracks (black bars) are shown for each biological replicate individually.

control vectors, expanded these cells and sorted $CD34^+GFP^+$ cells before cross-linking.

Three independent replicate experiments were performed, each with different cord blood batches. Before deep sequencing we tested each sample for enrichment at known target loci by qPCR (Figure 6A). Following sequencing, we only considered peaks that were present in at least two independent ChIP-Seq samples, and when the adjusted p-value was below 0.05. We detected only few CBX7 peaks in control samples and these were widespread across the genome. A substantial number of these peaks were not associated with promoter regions (classified by distances within 5kb of the transcriptional start site (TSS)). In contrast, peaks were strongly enriched at the TSS upon overexpression of *CBX7*, indicating that CBX7 acted specifically at core promoter regions. *CBX7* overexpression did not significantly alter the genome-wide occupancy of H3K27me3 and H3K9me3 in relation to TSSs (Figure 6B).

We next searched genome-wide for loci that were targeted by CBX7 and in addition covered by H3K9me3 or H3K27me3. We observed that 23 % of all CBX7 peaks were overlapping with H3K9me3, while 44 % were overlapping with H3K27me3. Furthermore, 13 % of all CBX7 peaks were associated with both H3K9me3 and H3K27me3 (Figure 6C), accumulating in ~1/3 of all CBX7 peaks being associated with H3K9me3. In agreement with our data, analysis of published ENCODE datasets of H3K9me3 and H3K27me3 ChIP-seq experiments in hematopoietic cell lines (K562 and GM12878) also showed partial overlap of genes decorated with H3K27me3 and H3K9me3 peaks (data not shown).

We then asked whether CBX7 and H3K9me3 were co-localized around TSSs, which would provide further evidence of a joint gene regulatory function. Indeed, in *CBX7* overexpressing $CD34^+$ cells ~20% of all TSSs marked with CBX7 were also marked with H3K9me3 (Figure 6E).

These molecular patterns are compatible with a model in which H3K9 methyltransferases act as binding partners of CBX7, at least for a subset of the genomic sites bound by CBX7. Interestingly, the genomic regions around TSSs of *CBX4* and *CBX8* loci were decorated with CBX7, as well as with H3K27me3 and H3K9me3, suggesting that *CBX7* antagonizes the two other *CBX* genes, whose expression resulted in a reduction in the CAFC-frequency upon overexpression (Figure 6D and Figure 1D).

To refine the list of direct targets of CBX7, we performed an integrative analysis of RNA-Seq and ChIP-Seq data and searched for genes that

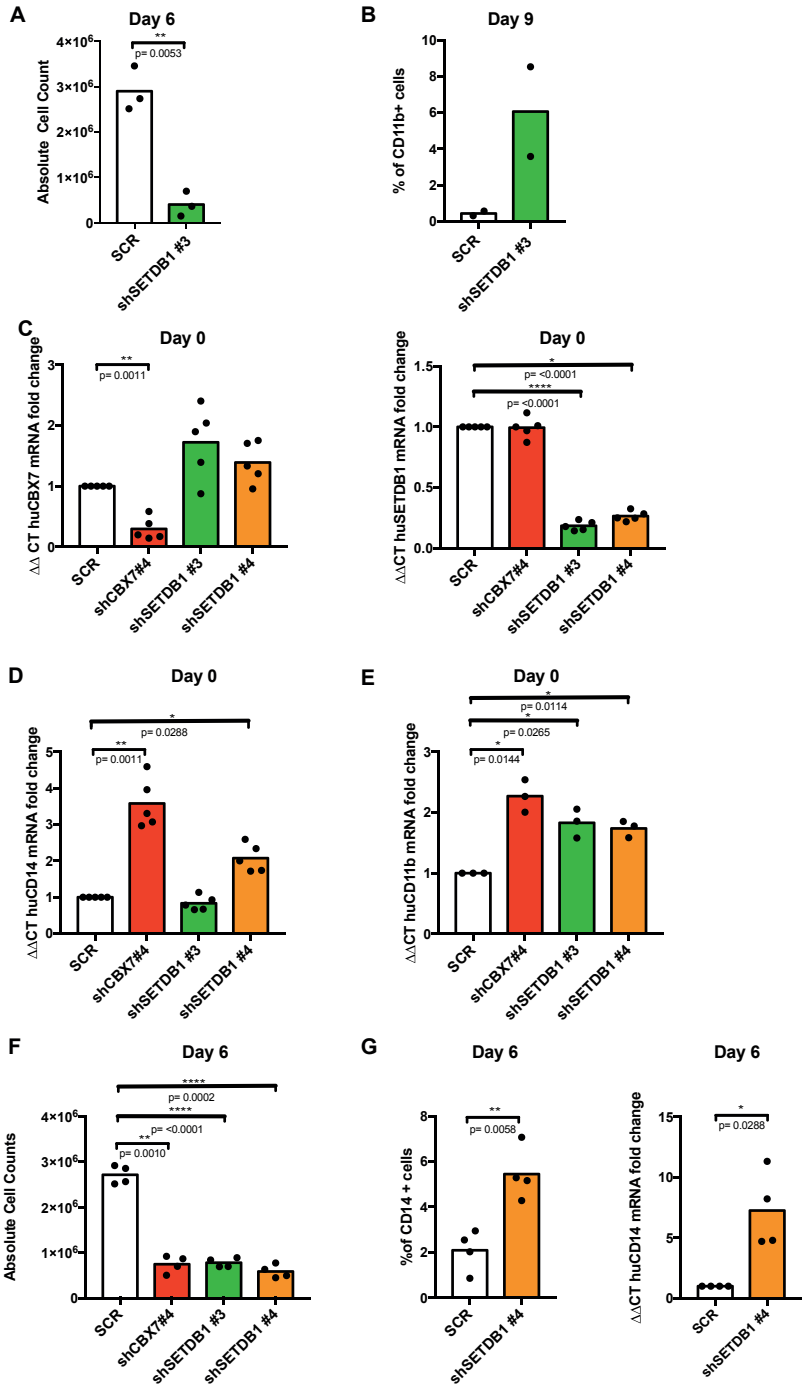
were downregulated upon overexpression of *CBX7* while at the same time covered with CBX7 ChIP-seq peaks. Out of 1183 repressed genes, 220 showed CBX7 peaks within 5kb around their transcriptional start site. Interestingly, an additional 63 genes were not marked within a region of 5kb around the TSS but instead around or within the gene body (Supplementary Figure 6A), indicating that CBX7 may also influence gene expression from regulatory regions outside the immediate vicinity of the TSS. All these 283 genes are likely to be primary targets of CBX7 in human HSPCs (Supplementary Table 1).

The large majority of these primary targets (246 out of the 283) were also marked by H3K27me3 within 5kb around the gene body, confirming the well know-interaction of CBX with the Polycomb repressive mark set by EZH2. Interestingly, and in agreement with our finding that CBX7 directly interacts with various H3K9 methyltransferases, 178 (i.e. 62%) of these direct CBX7 targets were also marked with H3K9me3 (Figure 6F). Most of those genes were marked by both H3K9me3 and H3K27me3, yet some repressed genes were exclusively marked with H3K9me3 (Figure 6F). Furthermore, we found 17 direct targets of CBX7 which showed increased H3K9me3 signal intensities upon overexpression of CBX7.

Collectively, these molecular signatures reveal functional non-canonical cross talk between Polycomb CBX proteins and H3K9 methylation, as first suggested by the physical interaction of CBX7 and SETDB1, EHMT1, and EHMT2.

SETDB1 and CBX7 share functional activity

As we identified the H3K9 methyltransferase SETDB1 as a novel CBX7 interacting protein, and as we found that approximately one third of the CBX7 genomic target loci were also covered by H3K9me3, we evaluated the function of SETDB1 in leukemic cells. Mutations in *SETDB1* are associated with the development of clonal hematopoiesis and *SETDB1* shows higher expression in leukemic stem cells compared to leukemic blasts in AML patient samples (Eppert et al., 2011; Steensma et al., 2015). Here we set out to investigate whether shRNA mediated knockdown of *SETDB1* in myeloid leukemic cells would phenocopy the effects observed upon *CBX7*-repression. Indeed, *SETDB1* knockdown strongly impaired proliferation of HL60 cells (Figure 7A). Strikingly, *SETDB1* knockdown resulted in increased expression of CD11b, similar as *CBX7* does in HL60


A
Day 6
Absolute Cell Count
4×10^6
3×10^6
2×10^6
1×10^6
0
**
p= 0.0053
SCR
shSETDB1 #3
B
Day 9
% of CD11b+ cells
10
8
6
4
2
0
SCR
shSETDB1 #3
C
Day 0
ΔΔ CT huCBX7 mRNA fold change
3
2
1
0
**
p= 0.0011
SCR
shCBX7#4
shSETDB1 #3
shSETDB1 #4
Day 0
ΔΔCT huSETDB1 mRNA fold change
1.5
1.0
0.5
0.0
*
p= <0.0001
****
p= <0.0001
SCR
shCBX7#4
shSETDB1 #3
shSETDB1 #4
D
Day 0
ΔΔCT huCD14 mRNA fold change
5
4
3
2
1
0
*
p= 0.0288
**
p= 0.0011
SCR
shCBX7#4
shSETDB1 #3
shSETDB1 #4
E
Day 0
ΔΔCT huCD11b mRNA fold change
3
2
1
0
*
p= 0.0114
*
p= 0.0265
*
p= 0.0144
SCR
shCBX7#4
shSETDB1 #3
shSETDB1 #4
F
Day 6
Absolute Cell Counts
4×10^6
3×10^6
2×10^6
1×10^6
0
****
p= 0.0002
****
**
p= <0.0001
**
p= 0.0010
SCR
shCBX7#4
shSETDB1 #3
shSETDB1 #4
G
Day 6
%of CD14 + cells
8
6
4
2
0
**
p= 0.0058
SCR
shSETDB1 #4
Day 6
ΔΔCT huCD14 mRNA fold change
15
10
5
0
*
p= 0.0288
SCR
shSETDB1 #4

cells (Figure 7B). Furthermore, *SETDB1* knockdown increased expression of CD11b in OCI-AML3 cells.

Also, similar to *CBX7*, knockdown of *SETDB1* in OCI-AML3 cells reduced proliferation and induced increased *CD14* mRNA expression immediately after transduction as well as six days later on protein level (Figure 7C right panel, 7D, 7G). We next confirmed that the *CD14* locus is a direct CBX7 target, as in primary $CD34^+$ cells *CD14* is in the top three downregulated genes upon overexpression of *CBX7* and is marked with CBX7, H3K27me3 and H3K9me3 (Supplementary Figure 6B). Surprisingly, knockdown of *SETDB1* resulted in increased expression of *CBX7* suggesting the presence of feedback-loops (Figure 7C). These experimental data indicate that CBX7 and SETDB1 jointly repress genes that are important for differentiation of leukemic cells towards mature myeloid cells.

DISCUSSION

In this study, we identify CBX7 as a regulator of self-renewal in normal and leukemic hematopoietic cells. We describe the complex molecular architecture of CBX7-induced self-renewal and discover a novel, biologically relevant, non-canonical role for CBX7 as a binding partner of multiple H3K9 methyltransferases, including SETDB1.

Polycomb CBX proteins are key components of the PRC1 complex, where their function is believed to be essential for recruitment of PRC1

Figure 7:

Repression of SETDB1 phenocopies repression of CBX7 in leukemic cells

(**A**) Absolute cell numbers after 6 days of culturing 150,000 HL60 cells upon short hairpin-mediated knockdown of SETDB1. (**B**) Percentage of CD11b+ HL60 cells after 9 days in culture upon knockdown of *SETDB1*. (**C**) Fold change of *CBX7* mRNA (left panel) and *huSETDB1* mRNA (right panel) expression in OCI-AML3 cells 24 hours after transduction with multiple short hairpins targeting *CBX7* or *SETDB1*. (**D+E**) Fold change of *CD11b* (**D**) and *CD14* (**E**) mRNA expression in OCI-AML3 cells 24 hours after transduction with multiple short hairpins targeting *CBX7* or *SETDB1*. (**F**) Absolute cell numbers after 6 days of culturing 150,000 OCI-AML3 cells upon short hairpin-mediated knockdown of *SETDB1* or *CBX7*. (**G**) Percentage of CD14+ cells (left panel) and fold change of *CD14* mRNA expression (right panel) upon knockdown of *SETDB1* in OCI-AML3 cells six days after sort. (All statistical analyses were performed using paired t-test, two-tailed)

to H3K27me3-modified genomic loci. Thus, the chromobox domain contained in all CBX proteins is able to recognize H3K27me3 modifications deposited by EZH1/2 as part of the PRC2 complex, which contributes to repression of target genes. Whereas the Drosophila genome contains a single *cbx* gene, during evolution amplification of *CBX* homologs has occurred in mammals. CBX2, -4, -6, -7, and -8 have all been described to be part of the PRC1 complex, and it is likely that various assemblies of PRC1 have distinct biological targets. In this project, we investigated the role of all five PRC1-CBX proteins in regulating human $CD34^+$ HSPCs. We show that CBX7 is uniquely able to enhance cell growth of primitive hematopoietic cell subsets. Additionally, transplantation of *CBX7* overexpressing $CD34^+$ cells resulted in enhanced long-term engraftment, multi-lineage differentiation potential, and an increased frequency of myeloid $CD33^+$ cells and primitive $CD34^+CD38^-$ cells in the bone marrow. These results are reminiscent of data of mouse Cbx7 which we reported earlier (Klauke et al., 2013), and establish CBX7 as an important evolutionary conserved regulator of self-renewal of human $CD34^+$ HSPCs.

Overexpression of *CBX7* resulted in repression of genes associated with differentiation and led to an upregulation of genes involved in cell cycle and DNA replication. ChIP-seq analysis showed that ~1/3 of the repressed differentiation-associated genes were direct CBX7 targets. Furthermore, many genes which were upregulated upon overexpression of *CBX7* are preferentially expressed by primitive hematopoietic cell subsets, and thus are likely to contribute to maintenance of the primitive phenotype. Overexpression of *CBX8* resulted in transcriptional consequences distinct from those observed after *CBX7* perturbation, illustrating that different CBX proteins regulate the expression of different genes, which likely explains the cell biological differences that we observed.

Our *in vitro*, as well as *in vivo*, data indicate that CBX7 regulates self-renewal activity of primitive cells. As we show that CBX7 represses genes important for differentiation, we hypothesized that CBX7 may also play a role in AML, where self-renewal is enhanced and conversely, differentiation is repressed. Here we show that knockdown of *CBX7* in leukemic cell lines affects their proliferation and results in derepression of genes that are normally expressed on differentiated cells.

The molecular mechanism by which CBX7 represses differentiation-inducing genes remains to be elucidated, but our studies strongly suggest that the interplay between the canonical, H3K27me3-mediated, and a newly

discovered, non-canonical, H3K9 mediated pathway plays an important role. Whereas the Drosophila Polycomb Cbx protein can only recognize H3K27me3 *in vitro* but not H3K9me3, biochemical studies have revealed that multiple mammalian CBX homologs can also bind to H3K9me3 in cell free systems, each with different binding affinities (Bernstein et al., 2006; Kaustov et al., 2011). So far, no H3K9 methyltransferases were described to interact with CBX proteins *in vivo*. As CBX proteins interact with trimethylated lysine residues on histone proteins via their chromodomain, we hypothesized that CBX proteins might also interact with non-histone proteins harboring a trimethylated lysine embedded in a motif highly similar to histone proteins. Indeed, our mass spectrometry analysis revealed multiple of such candidates. Interestingly, all four evolutionary conserved CBX interacting proteins (EHMT1 -a.k.a GLP-, EHMT2 a.k.a. G9A, SETDB1, and CDYL) have been shown to physically interact and are strongly associated with H3K9 methylation (Fritsch et al., 2010).

We focused our further studies on the interaction between CBX7 and SETDB1. SETDB1 is an H3K9 methyltransferase that is best known for its role in repressing the expression of endogenous retroviral elements in the genome (Collins et al., 2015). Interestingly, both SETDB1 and CBX7 have been identified as regulators of embryonic stem cell states (Bilodeau et al., 2009), but the role of SETDB1 in hematopoiesis has only recently emerged. Interestingly, mutations in *SETDB1* have been associated with clonal hematopoiesis in elderly individuals (Steensma et al., 2015). Recently, it has been shown that deletion of Setdb1 in murine hematopoietic stem cells results in bone marrow failure (Koide et al., 2016).

Biochemical studies have revealed that the chromodomain of CBX7 has high affinity for a trimethylated 24 amino-acid peptide, representing exactly the consensus amino-acid sequence of SETDB1 (amino acids 1157 to 1181). In fact, the affinity of CBX7 for this sequence is higher for peptides representing the amino-acid sequence of H3K27me3 or H3K9me3 (Kaustov et al., 2011). Three lysines residues of SETDB1 have been shown to be trimethylated (K490, K1170 and K1178), all could serve as putative binding sites for CBX7 (Hornbeck et al., 2015).

In accordance with the direct *in vivo* interaction between CBX7 and SETDB1, nearly one third of all CBX7 target loci were simultaneously covered with H3K9me3. In addition, the fact that 62 out of the 95 differentially expressed direct CBX7 target genes associated with differentiation were marked by both CBX7 and H3K9me3, strongly suggest that

self-renewal of human HSPCs is dependent on CBX7-mediated joint repression of target loci by methylation of both H3K27 and H3K9.

Reversely, we demonstrate that proliferation is decreased in leukemic cells when either *CBX7* or *SETDB1* is downregulated, or when CBX7 is pharmacologically inhibited. Similarly, it has been shown that exposure of murine AML blasts to the EHMT2 inhibitor UNC0638 leads to inhibition of growth and induction of myeloid differentiation (Lehnertz et al., 2014).

As for the exact molecular mechanism by which CBX7, SETDB1 and H3K9me3 interact, we hypothesize that such interactions are locus-specific and dependent on the exact composition of the protein complex involved. We propose that regulation follows a step-wise program, where trimethylated SETDB1 initially converts H3K9me or H3K9me2 into H3K9me3, resulting in attraction of PRC1 by binding of CBX7 to SETDB1. An alternative, not mutually exclusive possibility is that CBX7 first recognizes trimethylated SETDB1, by which it is then recruited to H3K9me2 loci to ensure further chromatin compaction. These recruitment models would be independent of H3K27me3/PRC2. At loci were both H3K27me3 and H3K9me3 histone marks are present, CBX7 could be recruited to both. It is interesting to note that one of the CBX7 binding proteins we identified, CDYL, can bind to EZH2 as well as to SETDB1 (Escamilla-Del-Arenal et al., 2013; Fritsch et al., 2010; Zhang et al., 2011) allowing for multiple alternative Polycomb and H3K9 methyltransferase interactions. Further elucidation of the daunting complexities by which seemingly independent epigenetic pathways converge will allow the understanding of the molecular machinery by which self-renewal is ensured. Disruption of these self-renewal pathways is likely to offer novel therapeutic opportunities in leukemia.

ACKNOWLEDGEMENTS

The authors wish to thank Geert Mesander, Johan Teunis and Theo Bijma of the Central Flowcytometry Unit of the University Medical Center Groningen, and Dick H Deckers and Jeroen Demmers of the Mass Spectrometry Facility of Erasmus MC for their excellent support. Furthermore, we would like to thank Kazemier I. for performing the MST measurements. This study was supported by a grant from the Dutch Cancer Society (RUG 2014-7178), the Netherlands Organization for Scientific Research (Mouse Clinic for Cancer and Aging) and a personal fellowship of the Deutsche Krebshilfe to JJ.

REFERENCES

Bagger, F. O., Sasivarevic, D., Sohi, S. H., Laursen, L. G., Pundhir, S., Sonderby, C. K., Winther, O., Rapin, N., and Porse, B. T. (2016). BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. Nucleic Acids Res 44, D917-924.

Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. Molecular and Cellular Biology 26, 2560-2569.

Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B., and Young, R. A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes & Development 23, 2484-2489.

Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 20, 1123-1136.

Cancer Genome Atlas Research, N., Ley, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., Hoadley, K., Triche, T. J., Jr., Laird, P. W., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368, 2059-2074.

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell 20, 845-854.

Collins, P. L., Kyle, K. E., Egawa, T., Shinkai, Y., and Oltz, E. M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. Proc Natl Acad Sci U S A 112, 8367-8372.

Comet, I., and Helin, K. (2014). Revolution in the Polycomb hierarchy. Nat Struct Mol Biol 21, 573-575.

Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poeppl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nat Med 17, 1086-1093.

Escamilla-Del-Arenal, M., da Rocha, S. T., Spruijt, C. G., Masui, O., Renaud, O., Smits, A. H., Margueron, R., Vermeulen, M., and Heard, E. (2013). Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. Mol Cell Biol 33, 5005-5020.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes & Development 17, 1870-1881.

Fritsch, L., Robin, P., Mathieu, J. R. R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A Subset of the Histone H3 Lysine 9 Methyltransferases Suv39h1, G9a, GLP, and SETDB1 Participate in a Multimeric Complex. Molecular Cell 37, 46-56.

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Research 43, D512-D520.

Kahn, T. G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., Reim, I., Pirrotta, V., and Schwartz, Y. B. (2016). Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. Nucleic Acids Res 44, 10132-10149.

Kamminga, L. M., Bystrykh, L. V., de Boer, A., Houwer, S., Douma, J., Weersing, E., Dontje, B., and de Haan, G. (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood 107, 2170-2179.

Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G. A., Li, Z., Vedadi, M., Schapira, M., et al. (2011). Recognition and specificity determinants of the human cbx chromodomains. J Biol Chem 286, 521-529.

Klauke, K., Radulovic, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., et al. (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat Cell Biol 15, 353-362.

Koide, S., Oshima, M., Takubo, K., Yamazaki, S., Nitta, E., Saraya, A., Aoyama, K., Kato, Y., Miyagi, S., Nakajima-Takagi, Y., et al. (2016). Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic

activation of nonhematopoietic genes. Blood 128, 638-649.

Koya, J., Kataoka, K., Sato, T., Bando, M., Kato, Y., Tsuruta-Kishino, T., Kobayashi, H., Narukawa, K., Miyoshi, H., Shirahige, K., and Kurokawa, M. (2016). DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation. Nat Commun 7.

Laurenti, E., Doulatov, S., Zandi, S., Plumb, I., Chen, J., April, C., Fan, J. B., and Dick, J. E. (2013). The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat Immunol 14, 756-763.

Lehnertz, B., Pabst, C., Su, L., Miller, M., Liu, F., Yi, L., Zhang, R., Krosl, J., Yung, E., Kirschner, J., et al. (2014). The methyltransferase G9a regulates HoxA9-dependent transcription in AML. Genes Dev 28, 317-327.

Min, J., Zhang, Y., and Xu, R. M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev 17, 1823-1828.

Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A., and Di Croce, L. (2013). RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. Cell Rep 3, 60-69.

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S. A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell 10, 47-62.

Nikoloski, G., Langemeijer, S. M. C., Kuiper, R. P., Knops, R., Massop, M., Tonnissen, E. R. L. T. M., van der Heijden, A., Scheele, T. N., Vandenberghe, P., de Witte, T., et al. (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat Genet 42, 665-667.

Ren, C., Morohashi, K., Plotnikov, Alexander N., Jakoncic, J., Smith, Steven G., Li, J., Zeng, L., Rodriguez, Y., Stojanoff, V., Walsh, M., and Zhou, M.-M. (2015). Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain. Chemistry & Biology 22, 161-168.

Rizo, A., Dontje, B., Vellenga, E., de Haan, G., and Schuringa, J. J. (2008). Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. Blood 111, 2621-2630.

Steensma, D. P., Bejar, R., Jaiswal, S., Lindsley, R. C., Sekeres, M. A., Hasserjian, R. P., and Ebert, B. L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood 126, 9-16.

Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol 9, 1428-1435.

Tadokoro, Y., Ema, H., Okano, M., Li, E., and Nakauchi, H. (2007). De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. The Journal of Experimental Medicine 204, 715.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., et al. (2012). RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. Cell 148, 664-678.

van den Boom, V., Rozenveld-Geugien, M., Bonardi, F., Malanga, D., van Gosliga, D., Heijink, A. M., Viglietto, G., Morrone, G., Fusetti, F., Vellenga, E., and Schuringa, J. J. (2013). Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells. Blood 121, 2452-2461.

Verhaak, R. G., Wouters, B. J., Erpelinck, C. A., Abbas, S., Beverloo, H. B., Lugthart, S., Lowenberg, B., Delwel, R., and Valk, P. J. (2009). Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica 94, 131-134.

Zhang, Y., Yang, X., Gui, B., Xie, G., Zhang, D., Shang, Y., and Liang, J. (2011). Corepressor Protein CDYL Functions as a Molecular Bridge between Polycomb Repressor Complex 2 and Repressive Chromatin Mark Trimethylated Histone Lysine 27. Journal of Biological Chemistry 286, 42414-42425.

SUPPLEMENTARY TABLE 1

Direct targets of CBX7 +/- 5000 bps around TSS

KDM7A	WNT5B	GPAT3	B3GALNT1	GOLGA2P10
MMP25	UST	PDE5A	CD14	OTUD7B
GIPR	HES1	PRDM5	TRH	DPYSL2
VCAN	FRMD4B	LEF1	CHD7	NDFIP2
DSG2	PODXL2	FGD4	KRT19	CA2
ARAP2	EPB41L5	GAS2L3	ZNF318	ID2
HDAC9	IGFBP2	PIF1	COL24A1	BAHD1
PTGER3	EPAS1	IRF8	LPAR3	SLC18A2
BCAR1	NID1	TOM1L1	SMPDL3A	ARHGAP42
NCKAP1	MYCL	RP5-862P8.2	NBEA	BMI1
LTK	CAMSAP2	CDC42BPA	RAPH1	NR1D2
SNX24	LGALSL	SOX13	NRIP3	FAM20C
NTN1	CNRIP1	RHOB	BNIP3	
ME1	MOB3B	CAMK2D	JUN	
PFKP	PLEKHG1	PLK2	MAF	
SPTB	DUSP4	PLA2G7	KCTD12	
MGAT4A	STX16	FBXO25	THBD	
ARHGAP10	CHST8	PARD3	FJX1	
VASH1	CDKN1A	HTR7	ZADH2	
TRIB2	PSD4	PTPRJ	FZD2	
PTPN18	PODXL	ST14	SGSH	
PDE8A	IL13RA1	CRIM1	EXT1	
FERMT2	LGALS3	DLG5	AP1S2	
ZNF532	PTPRE	SLC25A4	GAS6	
ARHGEF10L	RAP1GAP2	DST	KCNQ3	
RASAL2	WASF3	PTPN14	ZDHHC23	
CTTNBP2	CCNA1	TRIM36	MYBL1	
CRYBG3	MORC4	BCL2L11	MITF	
MAGI3	ARNTL	RBMS1	SHTN1	
OSTM1	SORT1	PID1	SHISA7	
SYDE2	PTGFRN	MCOLN2	RP11-43F13.1	
NRP1	FST	ANKH	ARL4C	
H2AFY2	DTNA	PRKCA	SPATS2L	
TPTEP1	SLC37A2	ZFYVE9	NTNG2	
KCNK10	GOLM1	FZD1	SULF2	
COCH	PSAT1	KIT	MYO6	
PPP1R16B	KCNMB4	FMNL2	IGF2R	
RGCC	NPL	AGAP1	MYO5A	
DGKH	USP44	CXCL16	DPP4	
CTSH	SLC41A2	ADCY9	S100A10	
CD276	ADAMTS7	PROK2	DNM3	
FZD3	TM6SF1	HSPA4L	TUSC1	
PTPRS	KLF4	ANXA5	UNC13B	
CPVL	TMOD1	NFIL3	PPP1R14C	
UNC5B	CTSV	OTUD1	COLGALT2	
TBC1D12	IRF4	UBTD1	MAFB	
RASD1	SLC22A23	FBN1	ZDBF2	
TBC1D9	THBS1	TRIM44	LGR4	
GLRB	PAQR5	PBK	AC104655.3	
DTX4	UACA	RAB31	ANKRD18B	
SLC15A3	TTLL7	STXBP6	FOXD2-AS1	
PRDM4	RASGEF1B	PCSK9	TMEM150C	

63 genes direct targets with peak in genebody +/- 5000

POU2F2	FBN2
COL23A1	PHLDA1
PDE4A	TACC1
RPS6KA2	EPS8
CHFR	CAST
COBLL1	DDAH1
C20orf194	UCHL1
CACNA1I	TIAM1
JPH1	MYO1E
BNIP3L	AUTS2
TMEM176B	KALRN
PMP22	NFIA
UNC93B1	FCMR
PHACTR1	SFMBT1
IL1R1	ARHGEF3
SGK1	SEPT8
RARRES1	DLC1
PLBD1	GPRC5C
EGR2	CLCN5
RUNX2	GATM
HSPA2	MYO1D
TNFRSF19	CACNB4
FOXJ1	CNR2
SH3BP5	FAM169A
KCNC3	MFAP3L
DSC2	TBC1D8
CEP162	SLFN12L
PKIB	RP13-104F24.2
CPM	PLXNA4
TRPC6	MIR34A
PLCB2	FRG1HP
CYP1B1	

MATERIAL AND METHODS

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Blocking agents for excluding unspecific binding:		
Human BD Fc block™	BD Bioscience	564220
CD16/ CD32 Mouse BD Fc block™	BD Bioscience	553142
Antibodies used for staining of differentiated hematopoietic cells		
Mouse anti human CD3 APC-Cy7 (SK7)	BD Bioscience	557832
Mouse anti human CD4 PE-Cy7 (SK3)	BD Bioscience	557852
Mouse anti human CD33 BV421(WM53)	BD Bioscience	562854
Mouse anti human CD45 APC (HI30)	BD Bioscience	560973
Mouse anti human CD19 PE (HIB19)	BD Bioscience	555413
BD Horizon Brilliant Stain buffer	BD Bioscience	563794
Antibodies used for staining of immature $CD34^+$ hematopoietic cells and LIN^+ cells		
Mouse anti human CD38 PE (HIT2)	BD Bioscience	555460
Mouse anti human CD90 AF700 (5E10)	BioLegend	328120
Mouse anti human CD34 APC (581)	BD Bioscience	555824
Mouse anti human CD45 RA PE-Cy7 (L48)	BD Bioscience	337186
Mouse anti human CD45 BV421 (HI30)	BioLegend	304032
Mouse anti human CD2 PE-Cy5 (RPA-2.10)	BioLegend	300210
Mouse anti human CD3 PE-Cy5 (UCHT1)	BioLegend	300410
Mouse anti human CD4 PE-Cy5 (RPA-T4)	BioLegend	300510
Mouse anti human CD7 PE-Cy5 (6B7)	BioLegend	343110
Mouse anti human CD8 PE-Cy5 (RPA-T8)	BioLegend	301010
Mouse anti human CD19 PE-Cy5 (HIB19)	BioLegend	302210
Mouse anti human CD20 PE-Cy5 (2H7)	BioLegend	302308
Mouse anti human CD235a PE-Cy5 (HIR2)	BioLegend	306606
Mouse anti human CD11b PE-Cy5 (ICRF44)	BioLegend	301308
Mouse anti human CD14 PE-Cy5 (TuK4), TRI-COLOR	ThermoFisher Scientific	MHCD1406
Mouse anti human CD56 PE-Cy5 (MEM-188)	BioLegend	304608
Antibodies used for staining of OCI-AML3/HL60 cells for *in vitro* experiments		
Mouse Anti-Human Alexa Fluor® 700 CD14 Clone M5E2 (RUO)	BD Bioscience	557923
Mouse Anti-Human BV421 CD11b/MAC-1 (RUO)	BD Bioscience	562632
Antibodies used for staining of $CD34^+$ cells for *in vitro* experiments		
Human anti CD34 PE-Cy7 (8G12)	BD Bioscience	348811
Materials for Isolation of $CD34^+$ cells from Cord blood		
$CD34^+$ MicrobeadKit	Miltenyi Biotec	130-056-702
Lymphoprep	Stem cell technologies	7861
LS Columns	Miltenyi Biotec	130-042-401
Antibodies used for Chip-Seq		
H3K9me3 polyclonal antibody-Premium	Diagenode	C15410193
Anti H3K27me3	Merck	07-449
Anti CBX7	Merck	07-981
Pierce Protein A/G Magnetic beads	Thermo Scientific	#88803
Antibodies used for Western-Blot and Proximity Ligation Assay		
Monoclonal ANTI-FLAG® M2 antibody	Sigma	F3165-2 mg
Polyclonal rabbit anti human/mouse CBX7 p15	Santa Cruz Biotechnology	SC 70-232
SETDB1 Antibody (5H6A12)	Pierce Protein	MA5-15722
Medium		
IMDM 2% FCS	StemCell Technologies	# 07700
MethoCult™ H4435 Enriched	StemCell Technologies	#04435 and 04445
Myelocult™ H5100	StemCell Technologies	#05100 and 05150

REAGENT or RESOURCE	SOURCE	IDENTIFIER
StemSpan™ SFEM	StemCell Technologies	# 09650
Source of consensus cDNA		
cDNA of huCBX2 (NM_005189), transcript variant 1	Origene	SC303599
cDNA of huCBX4 (NM_003655)	Origene	SC117841
cDNA of huCBX6	ThermoFisher	MHS6278-202759205 (now available via Dharmacon)
cDNA of huCBX7	ThermoFisher	MHS6278-202760094 (now available via Dharmacon) (site-directed mutagenesis to consensus cDNA, base exchange G230A and C715G)
cDNA of huCBX8	Kind gift from K. Hansen, Kopenhagen	(Dietrich et al., 2007)
cDNA of huSETDB1	Kind gift from Frank Rauscher, The Wistar Institute, Philadelphia, USA and Lingwen Ding, Cancer Science Institute of Singapore Republic of Singapore	Rauscher FJ 3rd et al, Genes & Development 16(8): 919-32, April 15 2002.
Sequences of short hairpins		
pLKO.1_GFP_SCR	CAACAAGATGAAGAGCACCAA	A kind gift from Vincent van den Boom
pLKO.1_GFP_shCBX7 (#1)	CGGAAGGGTAAAGTCGAGTAT	A kind gift from Vincent van den Boom
pLKO.1_mCherry_SCR	CAACAAGATGAAGAGCACCAA	
pLKO.1_mCherry_shCBX7	GCTTCTGTGAGGTGGTTTAGC	
pLKO.1_mCherry_shSETDB1#3	GCTCAGATGATAACTTCTGTA	Was subloned from Sigma shSETDB1 (TRCN 276169)
pLKO.1_mCherry_shSETDB1#4	AGTTAGAGACATGGGTAATA	Was subloned from Sigma shSETDB1 (TRCN 276105)
Primer sequences for cloning		
GFP_MluI_fw	CGCAATTGATGGTGAGCAAGGGCGAGGA	
FLAG_MluI_rv	CGCAATTGCTTGTCATCGTCGTCCTTGTAGT	
huCBX7_Not1_fw	CGG CGG CCG CATGGAGCTGTCAGCCATCGGC	
FLAG_huCBX7_Not1_fw	CGG CGG CCG C ATG GAC TAC AAG GAC GAC GAT GAC AAG ATG GAG CTG TCA GCC ATC GGC	
huCBX7_Sal1_rv	CGGTCGACTCAGAACTTCCCACTGCGGTCT	
huCBX2_Not1_fw	CGG CGG CCG C ATGGAGGAGCTGAGCAGCGTG	
huCBX2_Sal1_rv	CGGTCGAC TCAGTAATGCCTCAGGTTGAAG	
huCBX4_Not1_fw	CGG CGG CCG C ATGGAGCTGCCAGCTGTTGG	
huCBX4_Sal1_rv	CGGTCGAC CTACACCGTCACGTACTCCTTG	
FLAG_huCBX4_Not1_fw	CGG CGG CCG C ATG GAC TAC AAG GAC GAC GAT GAC AAG ATGGAGCTGCCAGCTGTTGG	
huCBX6_Not1_fw	CGG CGG CCG C ATG GAC TAC AAG GAC GAC GAT GAC AAG ATGGAGCTGCCAGCTGTTGG	

REAGENT or RESOURCE	SOURCE	IDENTIFIER
huCBX6_Sal1_rv	CGGTCGAC TCACTTGCTCGC-CCCAATGC	
FLAG_huCBX8_Not1_fw	CGG CGG CCG C ATG GAC TAC AAG GAC GAC GAT GAC AAG ATGGAGCTTTCAGCGGTG	
FLAG_huSETDB1_Mlu1_fw	CGA CGC GTA TGG ACT ACA AGG ACG ACG ATG ACA AGA TGT CTT CCC TTC CTG GGT G	
huSETDB1_Xba_rv	CGTCTAGACTAAAGAAGAC-GTCCTCTGCATTCA	
Primer sequences for measuring gene expression analysis by qPCR		
hubetaAktin_fw	TCCCTGGAGAAGAGCTACGA	
Hubeta_Aktin_rv	AGCACTGTGTTGGCGTACAG	
huHPRT_fw	GAACGTCTTGCTCGAGATGTG	
huHPRT_rv	TCCAGCAGGTCAGCAAAGAAT	
huCBX7_fw	GCGGAAGGGTAAAGTCGAGT	
huCBX7_rv	ACCTCTCTTCCTATACCCCGA	
huCBX8_fw	TGGTCGCAGAAGTACAGCAC	
huCBX8_rv	CACGCTTTTTGGGGCCATAG	
huSETDB1_fw	CCAAATATGGGTGCTGTG AGGA	
huSETDB1_rv	TTC CAC TGG CTT GAA CTG GG	
huCD11b_fw	ACT TGC AGT GAG AAC ACG TAT G	
huCD11b_rv	AGA GCC ATC AAT CAA GAA GGC	
huCD14_fw	ACTTGCACTTTCCAGCTTGC	(Haghparast et al., 2011)
huCD14_rv	GCCCAGTCCAGGATTGTCAG	
Chip-qPCR primers		
Alpha Sat_fw_Primer#1	CTGCACTACCTGAAGAGGAC	(Wang et al., 2013)
Alpha_Sat_rv_Primer#1	GATGGTTCAACACTGTTACA	
Alpha Sat_fw_Primer#2	AAGGTCAATGGCAGAAAAGAA	(Moralli et al., 2013)
Alpha_Sat_rv_Primer#2	CAACGAAGGCCACAAGATGTC	
INK4B_fw	ATCACGGAGCAATAAACCCAAC	(Kheradmand Kia et al., 2009)
INK4B_rv	CAAGAGAAACAGCGACCTAACC	
INK4A_fw	ACCAAGACTTCGCTGACC	(Kheradmand Kia et al., 2009)
INK4A_rv	CAAGGAGGACCATAATTCTACC	
RPL27_fw	TCCGGACGCAAAGCTGTCATCG	(Kheradmand Kia et al., 2009)
RPL27_rv	TCCGGACGCAAAGCTGTCATCG	
EIF4A2_fw	TTTTGTAGCTGACCGAAGCA	(Takayama et al., 2014)
EIF4A2_rv	GCGCCCTATGACCTTCACTA	
Chemicals, Peptides and recombinant protein		
Hydrocortisone	StemCell Technologies	#07904
Human recombinant TPO	RnD	288-TP-200
Human recombinant FLT3L	RnD	308-FK-025
Human recombinant SCF	RnD	255-SC-050
Human recombinant IL3	Sigma	I1646
Mouse recombinant IL3	RnDSystems	403-ml-50
RetroNectin® 2,5 mg	Westburg/ Takara	T100/B
Hexadimethrine bromide	Sigma	h-9268
FuGene HD transfection reagent	Promega	E2312
Formaldehyde, 37% formaldehyde solution	Santa Cruz Biotechnology	Sc-203049
SDS Solution, 20%	Fisher Scientific	BP 1311-1
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Merck	#000000011873580001
QIAquick PCR Purification Kit	Quiagen	#28104

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Novagen Benzoase Nuclease, Purity 99%	Merck Millipore	70664
SnakeSkin Dialysis Tubing, 7K MWCO, 22mm	Thermo Scientific	68700
3xFlag® peptide	Sigma	F3290-4MG
CryoStor® CS10	StemCell Technologies	#07930
Protease Inhibitor Cocktail	Sigma	P8340
Anti-Flag M2 Magnetic beads	Sigma	M8823
NT Protein Labelling kit Red-NHS	Nanotemper	MO-L001
MS37452	Sigma	SML1405
Experimental Models: Cell Lines		
MS5 cells	DSMZ	ACC 441
HL60 cells	ATCC	CCL-240
	DSMZ	ACC 3
OCI-AML3 cells	DSMZ	ACC 582
32D cells		Kind gift from Ivo Touw
PG13 cells	ATCC	CRL-10686
Phoenix-ECO cells	ATCC	CRL-3214
K562 cells	ATCC	CCL-243
293FT	ThermoFisher Scientific	R70007
Experimental Models: Organisms/ Strains		
NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ		Mice were purchased from Charles River Laboratory (L'Arbesle Cedex, France) and bred in house.
Software and Algorithms		
Graphpad Prism (v5-7)	Graphpad Prism	https://www.graph-pad.com
ELDA	Hu, Y, and Smyth, GK (2009). ELDA: Extreme limiting dilution analysis for comparing depleted and enriched pop-ulations in stem cell and other assays. Journal of Immunological Methods 347, 70-78.	http://bioinf.wehi.edu.au/software/elda/
FlowJo	Version X.0.7	

Method details:

Cell Culture:

Phoenix-ECO cells were culture in DMEM + 1% P/S + 10% heat-inactivated FCS. PG13 cells were culture in DMEM + 1% P/S + 10% heat-inactivated FCS. HEK293FT cells were culture in DMEM +1% P/S + 10% heat-inactivated FCS. HL60 cells were cultured in RPMI+ 1% P/S+ 20% heat-inactivated FCS. OCI-AML3 cells were cultured in RPMI + 1%P/S + 10% heat-inactivated FCS. K562 cells were cultured in RPMI + 1% P/S + 10% heat-inactivated FCS.

Cloning of retroviral vector constructs:

The consensus cDNA of CBX2,4,6,7 and 8 and FLAG-tagged versions of the cDNA were inserted in the retroviral vector backbone of SF91-IRES-GFP (Klauke et al., 2013) upstream of IRES by PCR based cloning using Not1 and Sal1 restriction sites. Primers used for PCR based cloning are listed above. FLAG-tagged GFP vector was cloned by vector-PCR of SF91-FLAG tagged muCbx7 (Klauke et al., 2013) with MluI restriction site containing primers and subsequent ligation.

Cloning of FLAG-tagged huSETDB1 cDNA in pRRLA:

A FLAG-tagged versions of SETDB1 cDNA was inserted in the lentiviral vector backbone of pRRLA IRES-GFP upstream of IRES by PCR based cloning using Mlu1 and Xba1 restriction sites. Primers used for PCR based cloning are listed above.

Cloning of short-hairpins in lentiviral expression vectors:

Corresponding oligos for SCR, shCBX7, shSETDB1#3 and shSETDB1#4 were annealed and cloned into the empty pLKO.1_mCherry vector upon digestion with Age1 and EcoR1.

CD34^{+} cord blood isolation:

Cord blood was obtained from healthy full-term pregnancies after informed consent in accordance with the Declaration of Helsinki from the obstetrics department at the Isala Hospital in Zwolle, the Netherlands. Initially, cord blood volume and cell counts were measured and then diluted 1:1 with PBS+ 2 mM EDTA+0.5% BSA. Maximum 30 ml of diluted cord blood was carefully layered on 15 ml of Lymphoprep™ in a 50 ml falcon tube and centrifuged for 20 minutes, 800g, without brakes. Middle layer containing mononuclear cells was harvested and diluted 1:1 with PBS 2 mM EDTA 0.5% BSA and then centrifuged for 5 minutes at 800g. Cell pellets were collected and washed with PBS 2mM EDTA 0.5% BSA and centrifuged for 10 minutes at 200g. Immunomagnetic labeling and separation were performed according to the manufacturer's manual of the CD34 MicroBead Kit, human (Miltenyi Biotec). Cells were either used immediately for experiments or frozen in Cryostor CS10.

Transduction of 32D cells:

Initially, 300,000 Phoenix-ECO cells/well (of a six-well plate) were seeded in DMEM + 1%P/S + 10% FCS. On day 2 cells were transfected with 1 µg of plasmid

with the help of FuGene® in a 1:3 ratio. 24 hours after transfection medium was changed to RPMI+10%FCS+1%P/S. On day 4 non-treated six- well plates were coated with RetroNectin® according to the manufacturer's manual. Viral supernatant was harvested and filtered through a sterile syringe filter with a 0.45 µm pore size hydrophilic PVDF membrane. Then 2 ml of viral supernatant, 300,000 32D cells and hexadimethrine bromide (2 µg/ml) and muIL3 (10 ng/ml) were added/well. Six-well plates were centrifuged for 45 minutes at room temperature for 45 minutes at 400 g . 24 hours after transduction virus-supernatant was replaced by RPMI+10%FCS+1%P/S+ muIL3 (10 ng/ml).

Production of a stable retroviral producer cell line (PG13):

Initially, 300,000 Phoenix-ECO cells/well (of a six-well plate) were seeded in DMEM + 1%P/S + 10% FCS. On day 2 cells were transfected with 1 µg of plasmid with the help of FuGene® in a 1:3 ratio. 24 hours after transfection medium was refreshed and 10,000 PG13 cells were plated out/ well in a tissue culture treated six-well plate. 48 hours after transfection viral supernatant was harvested and used to transduce the retroviral packaging cell line PG13 with the help of Hexadimethrine bromide (2 µg/ml). One day after transduction medium of transduced PG13 was changed to DMEM+1%P/S+10%FCS and cultured at 37°C/5% CO2.

Production of lentiviral supernatant:

2.75×10^6 293FT cells were plated in gelatin coated cell-culture treated 10 cm dishes in DMEM+10%FCS+1%P/S and incubated overnight at 37°C/5%CO_2. On the next day cells were transfected with 3 µg of the pLKO.1 or pRRLA vector, 3 µg of the packaging plasmid pCMV8.91 and 0.7 µg of envelope plasmid VSV-G and 21 µl of FuGene®. On the next day medium was changed to either StemSpan™ SFEM or RPMI. Two days after transfection the virus was collected, filtered through a sterile syringe filter with a 0.45 µm pore size hydrophilic PVDF membrane and used either immediately for transduction or was frozen.

Retroviral virus production and transduction of $CD34^+$ cells:

24 hours before the first transduction round $CD34^+$ cells were prestimulated in StemSpan™ SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml at 37°C and 5% CO2. Medium of transduced PG13 cells was changed to StemSpan™ SFEM. On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin® according to the manufacturer's manual.

Then viral supernatant of virus-producing PG13 cells was harvested and filtered through a sterile syringe filter with a 0.45 µm pore size hydrophilic PVDF membrane. Between 500,000 and 1,000,000 $CD34^+$ cells were transduced with 2 ml of viral supernatant in the presence of SCF 100 ng/ml, FLT3L 100 ng/ml, TPO 100 ng/ml and Hexadimethrine bromide to a final concentration of 2 µg/ml. Six-well plates were centrifuged at 400g for 1 hour at room temperature. Transduction was repeated two times in 8-12 hour time intervals. After last transduction round medium was changed to StemSpan™ SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml.

Lentiviral transduction of $CD34^+$ cells:

$CD34^+$ cells were cultured in StemSpan™ SFEM with SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml 24 hours before first transduction round at 37°C and 5% CO_2. On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin according to the manufacturer's manual. Lentiviral supernatant was thawn on ice. Between 500,000 and 1,000,000 $CD34^+$ cells were transduced with 2 ml of viral supernatant in the presence of SCF 100 ng/ml, FLT3L 100 ng/ml, TPO 100 ng/ml and Hexadimethrine bromide 2 µg/ml. Six-well plates were centrifuged at 400g for 1 hour at room temperature. Transduction was repeated once in 8-12 hour time intervals. After last transduction round medium was changed back to StemSpan™ SFEM containing SCF 100 ng/ml, FLT3L 100 ng/ml and TPO 100 ng/ml.

Lentiviral transduction of HL60 and OCI-AML3 cells:

On the day of transduction not tissue-cultured six-well plates were coated with RetroNectin according to the manufacturer's manual. Between 300,000 and 500,000 cells were transduced in 2 ml of viral supernatant containing Hexadimethrine bromide 2 µg/ml. Six-well plates were centrifuged at 400G for 1 hour at room temperature. Transduction was repeated once in 8–12 hour time intervals. After last transduction round medium was changed back to RPMI+1%P/S+10% (OCI-AML3) or 20% of FCS (HL60). At several time points cells were counted manually with a hemacytometer.

MS37452 treatment of OCI-AML3 cells:

Initially, 500.000 OCI-AML3 cells/well (of a six-well plate) were seeded in RPMI+1%P/S+10% heat-inactivated FCS supplemented with MS37452 (dissolved in DMSO at a concentration of 50µM) at different concentrations. After four days cells were counted manually using a hemacytometer.

FACS analysis of HL60 and OCI-AML3 cells:

Samples were incubated with Human BD Fc block™ to prevent unspecific binding at 4°C in the dark. After blocking, 5 µl mouse anti huCD11b BV421 and/or 5 µl mouse anti huCD14 Alexa Fluor 700 antibodies were added and samples were incubated for 20-25 minutes at 4°C in the dark. Afterwards, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II.

Sort of GFP^+CD34^+ cells (MoFlo Astrios and MoFloXDP):

24 hours after the last transduction round cells were harvested, washed and re-suspended in PBS+BSA 0.2%. Cells were incubated with Human BD Fc block™ for 15 minutes according to the manufacturer's manual to prevent unspecific binding. After blocking, 5 µl of mouse anti huCD34 PE-Cy7 was added and incubated at 4°C for 20 minutes. Cells were washed with PBS + BSA 0.2% and resuspended in PBS +BSA 0.2% with the viability dye (PI).

Colony-forming unit assay:

All experiments with $CD34^+$ cells were performed with single (not pooled) cords (except Chip-Seq experiments). 5,000 $CD34^+GFP^+$ cells were sorted in a FACS tube containing 1 ml of IMDM 2% FBS™. 0.3 ml of the sorted cells in IMDM 2%FCS™ were added to a pre-aliquoted 3 ml MethoCult™ tube. Afterwards the tube was vortexed for at least 4 seconds and then let stand for a minimum of 5 minutes. For dispensing the MethoCult™ mixture into 35 mm culture dishes a 3 ml syringe with a 16 gauge blunt-end needle was used to add 1.1 ml per dish. Dishes were cultured at 37°C / 5% CO2 conditions. Colonies were counted and typed after 14 days. For replating cells from primary dishes were harvested, centrifuged and counted with a hemocytometer. For CBX7 5,000 cells and for CBX8 30,000 cells were plated out as described above. Control cells were plated out at same cell numbers as experimental groups. Figure 1A+B show single data points of each experiment. Each single experiment was performed in duplicates and the average of the technical replicates was plotted.

Cobblestone area-forming cell assay:

96-well flat-bottom plates were pre-coated with 0.1% gelatin. Two days before sort 10,000 MS5 cells were plated in 200 µl of Myelocult™ H5100 supplemented with 10^{-6} M hydrocortisone + 1% P/S. On the next day cells were radiated with 30 Gy. One day post radiation $CD34^+GFP^+$ cells were sorted directly into 96-well plates at limiting dilution and cultured for 5 weeks with weekly performed

half-medium changes. Cobblestones were analyzed with a phase-contrast microscope. Frequency of each experiment was calculated with ELDA software. (Hu and Smyth, 2009) Figure 1D shows single data points of each experiment. (The Y-axis indicates the number of cells that need to be plated for a CAFC to develop.)

Long-term culture initiating cell assay:

96-well flat-bottom plates were pre-coated with 0.1% gelatin. Two days before sort 10,000 MS5 cells were plated out in 200 µl of Myelocult™ H5100 supplemented with 10^{-6} M hydrocortisone.

Cells were sorted directly into 96-well plates at limiting dilution and cultured for 5 weeks with weekly performed half-medium changes. After 5 weeks medium was replaced by Methocult H4335 and incubated for two further weeks at 37°C/5% CO2. Colony-formation was accessed by phase-contrast microscopy. Frequency of each experiment was calculated with ELDA software (Hu and Smyth, 2009).

(The Y-axis indicates the number of cells that need to be plated for a LTC-IC to develop)

Suspension culture experiment with *CBX7*, *CBX8* and *EV* overexpressing cells:

100,000 $CD34^+GFP^+$ cells were sorted in a six-well plate containing 2 ml of Myelocult™ H5100, 1% P/S, 10^{-6} M hydrocortisone, TPO 100 ng/ml, IL3 50 ng/ml, SCF 100 ng/ml, FLT3L 100 ng/ml. Cells were cultured at 37°C/5% CO2. Cells were counted manually with a hemacytometer every week and 100,000 cells were reseeded under the same conditions.

Xenotransplantation of transduced $CD34^+$ cells 24 hours after last transduction round:

Mouse experiments were performed in line with international and national guidelines. All experiments were approved by the Institutional Animal Care and Use Committee of the University of Groningen (IACUC-RUG).

For all xenotransplantation studies, we performed single cord transplantations of freshly isolated $CD34^+$ cord blood cells. Female 11-22 weeks old NOD. Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice were radiated three hours before transplantation with 1.8 Gy. In each experiment age of mice was balanced with maximum 2 weeks of difference between the experimental and control group. No antibiotic prophylaxes after radiation was given.

Isolation of $CD34^+$ cells and transduction was performed as described above. 24 hours after transduction the percentage of $CD34^+GFP^+$ cells was determined after Fc blocking and staining with $CD34^+$PE-Cy7 as described above. Cells were harvested and counted manually with a hemocytometer and trypan blue and resuspended in PBS. In total equivalents of 200,000 $CD34^+GFP^+$ cells were transplanted per mouse via retro-orbital injection. A small aliquot was kept in culture for determining the exact number of transplanted $CD34^+GFP^+$ cells 24 hours later. Because the transduction efficiency of experimental group was always lower than in the control group only absolute percentages of GFP engraftment are illustrated.

Xenotransplantation of transduced $CD34^+$ cells after one week in vitro culture:

Mouse experiments were performed in line with international and national guidelines and all experiments were approved by the Institutional Animal Care and Use Committee of the University of Groningen (IACUC-RUG).

For all xenotransplantation studies, we performed single cord transplantation of freshly isolated transduced $CD34^+$ cord blood cells. Female 10–20 weeks old NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mice were radiated three hours before transplant with 1.8 Gy. In the first experiment two control mice and one experimental mouse were 9 weeks older than 4 experimental mice. In the second experiment mice were between 15.5 and 19.5 weeks old upon transplant, age between both groups was balanced.

Isolation of $CD34^+$ cells was performed as described above. After isolation, cells were cultured in StemSpan™ SFEM exposed to FLT3L, TPO and SCF each with 100 ng/ml. Then three transduction rounds in a 24 hours time intervals were performed. 24 hours after the last transduction round cells were cultured for further 96 hours, so that cells were after isolation for one week *in vitro* cultured before transplantation.

On the day of transplantation percentage of $CD34^+GFP^+$ cells was determined after Fc blocking and staining with $CD34^+$PE-Cy7 as described above. Cells were harvested and counted manually with a hemocytometer and trypan blue and resuspended in PBS. In total equivalents of 1.5 million of $CD34^+GFP^+$ cells were transplanted per mouse via retro-orbital injection under general anesthesia. Because transduction efficiency of experimental group was always lower than in the control group only absolute percentages of GFP engraftment are illustrated.

Bleeding of xenotransplanted NSG mice:

Beginning 6-weeks after transplantation chimerism in peripheral blood was determined in 4-week intervals. Blood samples were taken under general anesthesia via retro-orbital bleeding. Blood was lysed with ammonium chloride, washed two times with PBS+BSA 0.2% and resuspended in 50 µl of PBS+BSA 0.2%. Samples were incubated with Human BD Fc block™ and CD16/CD32 Mouse BD Fc block™ to prevent unspecific binding. After incubation for 10 minutes at room temperature antibody master mix and BV stain buffer was added and samples were incubated for 20-25 minutes at 4°C in the dark. Afterwards, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II.

Bone marrow analysis of NSG mice:

Mice were sacrificed and dissected under general anesthesia after the end of the experiment or reaching the human endpoint of the experiment. Bones (Femur, tibia, fibula and pelvis) were collected and cleaned. Bones were crushed in the presence of PBS+0.2% PBS with a mortar and pestle. The obtained cell suspension was filtered through a 40 µm filter. Remaining erythrocytes were lysed with ammonium chloride. Cells were then pelleted by centrifugation and resuspended in PBS+BSA 0.2%. For preventing unspecific binding samples were incubated with Human BD Fc block™ and CD16/CD32 Mouse BD Fc block™.

After incubation for 10 minutes at room temperature, antibody master mix and BV stain (only for lineage-specific staining) buffer was added and samples were incubated for 20–25 minutes at 4°C in the dark. Afterwards, cells were washed and resuspended with PBS+BSA 0.2% containing a viability dye (PI). Samples were analyzed on a BD FACSCanto II (Lineage staining) and LSRII (progenitor staining).

RNA-Seq of $CD34^+$ cells:

100,000 GFP^+ $CD34^+$ cells were sorted into lysis buffer 4 days post-transduction. RNA was extracted using the Nucleospin RNA XS Kit, with the addition of a second elution step to increase yield. RNA quality and quantity was assessed using the Bioanalyzer RNA total Pico Assay. RNA samples with an RNA-integrity >8 were processed for RNA-seq library preparation using the SMARTer Stranded Total RNA-seq Kit. Briefly, 10 ng of total RNA was reverse transcribed using random primers and amplified via PCR during which barcoded Illumina adapters were added. After amplification, ribosomal RNA and mitochondrial cDNA were removed by annealing specific R-probes,

resulting in cleavage of ribosomal and mitochondrial cDNA in the presence of ZapR. After cleavage of ribosomal and mitochondrial cDNA, the remaining cDNA was amplified again during another round of PCR. The nM concentration of RNA-seq libraries were quantified based on library size (Bioanalyzer) and cDNA concentration (Qubit) and normalized to 2 nM prior to pooling. RNA-sequencing was performed on an Illumina HiSeq 2500 machine, three single-end runs with a read length of 63-64nt, resulting in fastq samples consisting of 26 to 64 million reads. Sample mapping was done with STAR (version 2.5.1b-2.5.2b), a custom genome index was build using Genecodegenes. org release 24 (GRCh38.p5) Ensembl 83, December 2015. STAR outputs read counts per gene, these where filtered by removing ribosomal and transfer RNA. Differential expression (DE) analyses by EdgeR (version 3.16.2), with upper quartile normalization. DE genes per condition ranked by p-value and adjusted for multiple testing using a Benjamini- Hochberg method. The final DE gene lists were filtered by FDR<0.05.

The RNA-seq data are deposited at ENA (PRJEB22831).

Gene Annotation and GO search was done using String database (https://string-db.org). Corresponding annotated and GO files were downloaded from the site and further analysed with custom scripts. GOChord plots were done in R (GOplot packages).

RNA expression data from Laurenti et al. (Laurenti et al., 2013), were obtained from GEO (GSE42141). Expression heatmap was done using R using heatmap.2 function in gplot library. Data were clustered by correlation and z-transformed. Gene Set-Enrichment Analysis (Subramanian, Tamayo, et al. 2005, PNAS 102, 15545-15550 and Mootha, Lindgren, et al. (2003, Nat Genet 34, 267-273)) was performed on a preranked list of all differentially expressed genes (FDR<0.1). The number of permutations was set to 1000, with exclusion of filtersets <10. (Furthermore, we applied the following filters scoring_sheme weighted, norm meandiv, make_sets true, gui false, set_max 500, set_min 15, npmerm 1000)

Chip-Seq of transduced CD34$^+$ cells:

Frozen human CD34$^+$ enriched cord blood cells were thawn, pooled (Batch 1= 21 cords, Batch 2= 21 cords, Batch 3= 31 cords) and cultured in StemSpan™ SFEM with TPO, SCF and FLT3L (each 100 ng/ml) for prestimulation. After 24–48 hours three transduction rounds with retroviral supernatant for either EV or CBX7 at 24-hour time intervals were performed. Transduction was performed as described above. After the final transduction round, cells were

further expanded in StemSpan™ SFEM with TPO, SCF and FLT3L (each 100 ng/ml). One week after thawing, cells were stained and $CD34^+GFP^+$ cells sorted as described above. Sorted cells were washed in ice-cold PBS+BSA 0,2%, centrifuged (450G, 5 min, 4°C) and resuspended in 1% cold formaldehyde for fixation. Tubes were incubated on a rotator at 4°C for 10 minutes. Fixation was stopped by adding glycine to a final concentration of 0.125 M. After addition of glycine cells were incubated on a rotator at 4°C for 5 minutes, washed two times with cold PBS, transferred to a low-adherent tube and resuspended in SDS buffer (NaCl 100mM, Tris-Cl pH8.1 50 mM, NaN_3 0.2%, 0.5 % SDS) + cOmplete™ protease inhibitors (1 tablet/ 50 ml SDS buffer). Samples were snapfrozen on dry ice and then transferred to -80°C for storage. The following table describes the number of transduced $CD34^+$ cells/ experiment and per antibody in million of cells.

	H3K9me3	H3K27me3	CBX7	IgG
1st Experiment	1.4	1.4	2	0.2
2nd Experiment	0.65	0.8	0.8	0.125
3rd Experiment	1	1	2	0.2

For chromatin-immunoprecipitation, samples were thawn, centrifuged for 5 minutes at 900 g at room temperature and the supernatant was discarded. Pellets were resuspended in 500 µl of IP buffer (30 ml SDS buffer + 15 ml Triton dilution buffer (100 mM Tris-Cl pH8.6 + 100 mM NaCl + 5 mM EDTA ph 8.0 + 0.2% NaN_3 + 5%Triton X-100 + cOmplete™ protease inhibitors) and sonicated to an average length of 400-500 bps (Bioruptor 30s on/30s off/cycle, high, in total 3 cycles). 5% of each sample was reversed crosslinked in TE with 1% SDS and 200 mM NaCl overnight at 65°C. On the next day decrosslinked DNA was isolated with the QIAquick PCR purification kit (Qiagen) and appropriate fragment length was confirmed via agarose gel electrophoresis. Protein A/G magnetic beads were washed three times with cold PBS and once with cold IP buffer. Crosslinked and fragmented samples were thawn and centrifuged for 30 min at 17,000g at 4°C. Samples were precleared by rotating at 4°C with 7.5 µl of washed beads for 1 hour. Samples were then distributed into several low adherence tubes for incubation with 5 µg of antibody and incubated overnight at 4°C on a rotating platform. The next day, samples were incubated with 20 µl of washed beads for 4 hours on a rotating platform at 4°C. Using a magnetic stand, the supernatant was removed and beads were washed 4x in a low salt buffer (150 mM) and once with TE. After the last washing step samples were

incubated in 1% SDS, 200 nM NaCl overnight at 65°C with 1100 RPM to reverse crosslinks. DNA was isolated with QIAquick PCR purification kit and enrichment of positive and negative loci was confirmed via qPCR.

ChIP libraries for sequencing were prepared with the Microplex Library Preparation Kit V2 (Diagenode, C05010012) according to the manufacturers protocol. The concentration of individual Chip-seq samples was determined based on library size (Agilent, Bioanalyzer 2100) and DNA concentration (Thermo Fisher, Qubit) and diluted to 2nM prior to pooling.

ChIP-Seq was performed on an Illumina NextSeq 500 machine, paired-end 79–80 bp,, a custom genome index was build using Genecodegenes.org release 25 (GRCh38.p7) Ensembl 85, July 2016. BAM mem (version 0.7.15) produced BAM files were generated and processed with Samtools (version 1.3.1). MACS2 (version 2.1.0) with the settings -f BAMPE –-nomodel –-broad –broad-cutoff -g 2.7e9 –kep-dup 1. The output BED files where analysed using bedtools (version 2.26.0) functions intersect, closest and Deeptools (version 2.4.2).

The Chip-seq data are deposited at ENA (PRJEB22344).

Mass spectrometry of pull-downs of FLAG-tagged-huCBX7, -huCBX8, -huCBX4 and -GFP in K562 cells and FLAG-tagged-muCbx7 and -GFP in 32D-cells:

32D cells and K562 cells were transduced as described above, sorted and expanded. Cells were harvested and washed. For each experiment 2-4 ml of cell pellets were used. Cell pellets were resuspended in 4-5 pellet volumes of ice-cold buffer A (10mM Hepes ph 7.6, 1.5mM MgCl2 and 10mM KCl, 0.5 mM DTT+ complete protease inhibitor), lysed for 10 minutes on ice and centrifuged for 10 min at 3000 rpm. Pellets were resuspended in 2 pellet volumes of buffer A. The cell suspension was homogenized with a Dounce Homogenizer with 10 strokes (pestle A). The homogenized suspension was centrifuged for 10 minutes with 3000 rpm. Supernatant was removed and pellets were centrifuged for 1 minute with 3000 rpm. Supernatant was removed and pellets were resuspended in 1.5 pellet volumes of buffer C (20 mM Hepes ph 7.6, 20% glycerol, 420 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM DTT, complete protease inhibitor). Nuclei suspension was homogenized with Dounce homogenizer with 10 strokes (pestle B). The suspension was then rotated on a rotor suspension at 4°C for 30 minutes and afterwards centrifuged for 15 min with maximal speed. Nuclear extracts were dialyzed with a SnakeSkin Dialysis tube to buffer D (20 mM HEPES [pH 7.6], 0.2 mM EDTA, 1.5 mM $MgCl_2$, 100 mM KCl, 20% glycerol, 0.5 mM DTT, 0.2 mM PMSF, 9.5 mg/l sodium Metabisulfite).

60 µl of anti-FLAG M2 agarose beads (Sigma) equilibrated and washed in buffer C-100 (20 mM Hepes ph 7.6, 20% glycerol, 1.5 mM MgCl2, 100 mM KCl, 0.2 mM EDTA, 0,02% NP40, 0.5 mM DTT, complete proteaseinhibitor) were added to 1.5 ml of nuclear extract in low-adherence microcentrifuge tubes and incubated for 3 hr at 4°C in the presence of 225 units of Benzonase (Novagen) on a rotator. Afterwards beads were washed 5-times with buffer C-100. Bound proteins were eluted four times with 60 µl of 3xFLAG-peptide solution (buffer c-100 + 0.2 mg/ml 3xFLAG peptide) for 15 minutes at 4°C. Efficiency of co-immunoprecipitation and elution was checked via Western blot with staining against FLAG. Most efficient elutions were pooled, TCA precipitated, and proteins separated by polacrylamide gel electrophoresis stained with (Invitrogen). MS was done in a label free format. For each identified peptide and protein spectral counts were calculated using commercial PEAK studio software, using standard filtering settings. Further, data were merged for all 10 slices into a single table, the sum of all spectral counts (per slice) was taken as a measure of the protein amount in the sample. Data were grouped into three experiments (Human, mouse 1 and mouse 2), sorted by spectral count abundance. The last 10% of the least abundant proteins were removed from each pull down list. The rest of the proteins were ranked in relative scale (0 for the most abundant to 0.9 as the least abundant). The cumulative abundance rank index was calculated for every candidate protein by subtracting each abundance rank index from the control rank index (GFP).

The list of proteins with Kme3 modifications was downloaded from Phosphosite database. Information for proteins abundance was downloaded from PaxDB site average abundance across all human samples was used for this analysis.

Cross comparison of the gene/protein lists from different databases was done using custom scripts. All illustrations were prepared in R and Python using standard graphic packages.

For identification of trimethylated interaction partners of CBX7 we only considered proteins which were in the top 20% of the relative ranked ordered CBX7-binding proteins and which have an abundance score below 100 (ppm) over all cell types and whose relative ranked score in the experimental sample is higher than in the control sample.

To confirm interaction of CBX7 and SETDB1 in K562 cells we performed Co-IP in FLAG-tagged CBX7 overexpressing K562 cells as described above and performed SDS Page. After protein transfer membrane was stained against CBX7 and SETDB1.

Detection of CBX7 and SETDB1 interaction in HL60 cells by DUOLINK in situ proximity ligation assay (PLA):

5×10^4 HL60 cells were fixed in ice-cold methanol for 5 minutes on cytospin slides. Interaction of the endogenous CBX7 and SETDB1 proteins in HL60 cells was assessed using the Duolink in situ Proximity Ligation Assay (PLA) (Olink Bioscience, Uppsala, Sweden), as described by the manufacturer. Used antibodies:

- Polyclonal rabbit anti human/mouse CBX7 p15 Santa Cruz Biotechnology, SC 70-232
- anti- SETDB1 Antibody (5H6A12) Pierce Protein, MA5-15722

Purification of FLAG-tagged protein:

Retroviral supernatant for overexpressing FLAG-tagged huCBX7 and lentiviral supernatant for overexpressing FLAG-tagged huSETDB1 was produced as described above. 293FT cells were transduced with two transduction rounds and expanded.

FLAG-tagged SETDB1 and CBX7 was purified using anti-Flag M2 magnetic beads (Sigma). Frozen pellets were resuspended in 10 vol (v/v) lysis buffer (50 mM Tris-HCl pH 7.5, 600 mM NaCl, 2 mM EDTA, 1.0% NP40 (v/v), 1:1000 µl Protease Inhibitor Mix (Sigma-Aldrich). Lysates were incubated for 30 min on ice while being inverted (3 times) every 5 min. After incubation the lysate was forced 10 times through a 26 g needle. Debris was pelleted by centrifugation (20000 g, 30 min, 4°C) and the supernatant transferred to a new tube. 1 original pellet vol. (v/v) anti-Flag M2 magnetic beads (Sigma-Aldrich) was washed 4 times with 1 ml TBS (50 mM Tris-Cl, pH 7.6; 150 mM NaCl) and 1 time with lysis buffer. The supernatant and the washed beads were combined and antigen capture was performed for 4 hr at 4°C using a head-over-tail rotator (HOT). After incubation, samples were placed on a magnetic stand for 1 min. The supernatant was removed and the bead-bound protein washed 3 times with 1ml lysis buffer (5 min HOT, 1 min magnetic stand) and 3 times with 1 ml wash buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM EDTA, 0.1% NP40 (v/v)). The FLAG-tagged protein was eluted by adding 150 µl elution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 3xFlag-Peptide 1mg/ml) and incubation for 5 hr at 4°C. 20 µl 85% glycerol were added to the eluted protein. Proteins were stored at -20°C.

Microscale thermophoresis (MST):

SETDB1 was labeled with Monolith NT Protein Labelling kit Red-NHS (Nanotemper). For binding reactions, a constant of 25 nM NT-647 labeled

SETDB1 was used and a 1:1 serial dilution of CBX7, with 25 nM as the highest concentration

Binding reactions were prepared in 1x MST buffer supplemented with 0,05% Tween-20 to a total volume of 10 µl. Microscale thermophoresis (MST) analysis was performed on a Monolith NT.115 (Nanotemper) with 40% LED and 20% MST power using standard treated capillaries (Nanotemper). Kd was calculated a using the MO affinity analysis software (Nanotemper).

REFERENCES

Bagger, F. O., Sasivarevic, D., Sohi, S. H., Laursen, L. G., Pundhir, S., Sonderby, C. K., Winther, O., Rapin, N., and Porse, B. T. (2016). BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. Nucleic Acids Res *44*, D917-924.

Bernstein, E., Duncan, E. M., Masui, O., Gil, J., Heard, E., and Allis, C. D. (2006). Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. Molecular and Cellular Biology *26*, 2560-2569.

Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B., and Young, R. A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes & Development *23*, 2484-2489.

Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev *20*, 1123-1136.

Cancer Genome Atlas Research, N., Ley, T. J., Miller, C., Ding, L., Raphael, B. J., Mungall, A. J., Robertson, A., Hoadley, K., Triche, T. J., Jr., Laird, P. W., *et al.* (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med *368*, 2059-2074.

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell *20*, 845-854.

Collins, P. L., Kyle, K. E., Egawa, T., Shinkai, Y., and Oltz, E. M. (2015). The histone methyltransferase SETDB1 represses endogenous and exogenous retroviruses in B lymphocytes. Proc Natl Acad Sci U S A *112*, 8367-8372.

Comet, I., and Helin, K. (2014). Revolution in the Polycomb hierarchy. Nat Struct Mol Biol *21*, 573-575.

Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poeppl, A., Ling, V., Beyene, J., *et al.* (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nat Med *17*, 1086-1093.

Escamilla-Del-Arenal, M., da Rocha, S. T., Spruijt, C. G., Masui, O., Renaud, O., Smits, A. H., Margueron, R., Vermeulen, M., and Heard, E. (2013). Cdyl, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. Mol Cell Biol *33*, 5005-5020.

Fischle, W., Wang, Y., Jacobs, S. A., Kim, Y., Allis, C. D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes & Development *17*, 1870-1881.

Fritsch, L., Robin, P., Mathieu, J. R. R., Souidi, M., Hinaux, H., Rougeulle, C., Harel-Bellan, A., Ameyar-Zazoua, M., and Ait-Si-Ali, S. (2010). A Subset of the Histone H3 Lysine 9 Methyltransferases Suv39h1, G9a, GLP, and SETDB1 Participate in a Multimeric Complex. Molecular Cell *37*, 46-56.

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Research *43*, D512-D520.

Kahn, T. G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., Reim, I., Pirrotta, V., and Schwartz, Y. B. (2016). Interdependence of

PRC1 and PRC2 for recruitment to Polycomb Response Elements. Nucleic Acids Res *44*, 10132-10149.

Kamminga, L. M., Bystrykh, L. V., de Boer, A., Houwer, S., Douma, J., Weersing, E., Dontje, B., and de Haan, G. (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. Blood *107*, 2170-2179.

Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., Wasney, G. A., Li, Z., Vedadi, M., Schapira, M., *et al.* (2011). Recognition and specificity determinants of the human cbx chromodomains. J Biol Chem *286*, 521-529.

Klauke, K., Radulovic, V., Broekhuis, M., Weersing, E., Zwart, E., Olthof, S., Ritsema, M., Bruggeman, S., Wu, X., Helin, K., *et al.* (2013). Polycomb Cbx family members mediate the balance between haematopoietic stem cell self-renewal and differentiation. Nat Cell Biol *15*, 353-362.

Koide, S., Oshima, M., Takubo, K., Yamazaki, S., Nitta, E., Saraya, A., Aoyama, K., Kato, Y., Miyagi, S., Nakajima-Takagi, Y., *et al.* (2016). Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. Blood *128*, 638-649.

Koya, J., Kataoka, K., Sato, T., Bando, M., Kato, Y., Tsuruta-Kishino, T., Kobayashi, H., Narukawa, K., Miyoshi, H., Shirahige, K., and Kurokawa, M. (2016). DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation. Nat Commun *7*.

Laurenti, E., Doulatov, S., Zandi, S., Plumb, I., Chen, J., April, C., Fan, J. B., and Dick, J. E. (2013). The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. Nat Immunol *14*, 756-763.

Lehnertz, B., Pabst, C., Su, L., Miller, M., Liu, F., Yi, L., Zhang, R., Krosl, J., Yung, E., Kirschner, J., *et al.* (2014). The methyltransferase G9a regulates HoxA9-dependent transcription in AML. Genes Dev *28*, 317-327.

Min, J., Zhang, Y., and Xu, R. M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev *17*, 1823-1828.

Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A., and Di Croce, L. (2013). RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. Cell Rep *3*, 60-69.

Morey, L., Pascual, G., Cozzuto, L., Roma, G., Wutz, A., Benitah, S. A., and Di Croce, L. (2012). Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell Stem Cell *10*, 47-62.

Nikoloski, G., Langemeijer, S. M. C., Kuiper, R. P., Knops, R., Massop, M., Tonnissen, E. R. L. T. M., van der Heijden, A., Scheele, T. N., Vandenberghe, P., de Witte, T., *et al.* (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat Genet *42*, 665-667.

Ren, C., Morohashi, K., Plotnikov, Alexander N., Jakoncic, J., Smith, Steven G., Li, J., Zeng, L., Rodriguez, Y., Stojanoff, V., Walsh, M., and Zhou, M.-M. (2015). Small-Molecule Modulators of Methyl-Lysine Binding for the CBX7 Chromodomain. Chemistry & Biology *22*, 161-168.

Rizo, A., Dontje, B., Vellenga, E., de Haan, G., and Schuringa, J. J. (2008). Long-term maintenance of human hematopoietic stem/progenitor cells by expression of BMI1. Blood *111*, 2621-2630.

Steensma, D. P., Bejar, R., Jaiswal, S., Lindsley, R. C., Sekeres, M. A., Hasserjian, R. P., and Ebert, B. L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. Blood *126*, 9-16.

Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

Tadokoro, Y., Ema, H., Okano, M., Li, E., and Nakauchi, H. (2007). De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. The Journal of Experimental Medicine *204*, 715.

Tavares, L., Dimitrova, E., Oxley, D., Webster, J., Poot, R., Demmers, J., Bezstarosti, K., Taylor, S., Ura, H., Koide, H., *et al.* (2012). RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. Cell *148*, 664-678.

van den Boom, V., Rozenveld-Geugien, M., Bonardi, F., Malanga, D., van Gosliga, D., Heijink, A.

M., Viglietto, G., Morrone, G., Fusetti, F., Vellenga, E., and Schuringa, J. J. (2013). Nonredundant and locus-specific gene repression functions of PRC1 paralog family members in human hematopoietic stem/progenitor cells. Blood *121*, 2452-2461.

Verhaak, R. G., Wouters, B. J., Erpelinck, C. A., Abbas, S., Beverloo, H. B., Lugthart, S., Lowenberg, B., Delwel, R., and Valk, P. J. (2009). Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica *94*, 131-134.

Zhang, Y., Yang, X., Gui, B., Xie, G., Zhang, D., Shang, Y., and Liang, J. (2011). Corepressor Protein CDYL Functions as a Molecular Bridge between Polycomb Repressor Complex 2 and Repressive Chromatin Mark Trimethylated Histone Lysine 27. Journal of Biological Chemistry *286*, 42414-42425.